ELSEVIER
ACADEMIC
PRESS

Encyclopedia of APPLIED PSYCHOLOGY



Editor-in-Chief
CHARLES SPIELBERGER

VOLUME 3 P2–Y

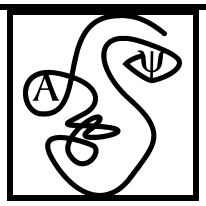


Editor–in–Chief

CHARLES D. SPIELBERGER

University of South Florida
Tampa, Florida, USA

Charles D. Spielberger is Distinguished Research Professor of Psychology and Director of the Center for Research in Behavioral Medicine and Health Psychology at the University of South Florida. Author, co-author, or editor of more than 350 professional publications, Spielberger's current research focuses on anxiety, anger, curiosity, depression, job stress, and lifestyle factors that contribute to the etiology and progression of cancer and cardiovascular disorders. His *State-Trait Anxiety Inventory* (1970, 1983), with translations in 66 languages and dialects, has become a standard international measure of anxiety. He has served as President of the International Association of Applied Psychology, the International Council of Psychologists, and the International Stress Management Association and was the 100th President of the American Psychological Association.



Editorial Advisors

John W. Adair
University of Manitoba
Winnipeg, Manitoba, Canada

Michael Harris Bond
Chinese University of Hong Kong
Hong Kong, China

Pieter J. D. Drenth
Vrije Universiteit Amsterdam
Amsterdam, The Netherlands

Edwin A. Fleishman
George Mason University
Potomac, Maryland, USA

Raymond D. Fowler
American Psychological Association
Washington, D.C., USA

Michael Frese
Justus-Liebig University of Giessen
Giessen, Germany

Machiko Fukuhara
Tokiwa University
Mito, Japan

James Georgas
University of Athens
Athens, Greece

Frank Heller
Tavistock Institute of Human Relations
London, United Kingdom

Stevan Hobfoll
Kent State University
Kent, Ohio, USA

Anna Leonova
Moscow State University
Moscow, Russian Federation

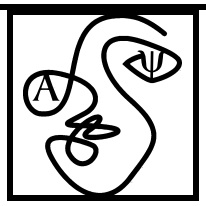
Claude Levy-Leboyer
Université René Descartes
Paris, France

Joe Matarazzo
Oregon Health Sciences University
Portland, Oregon, USA

Charles Spielberger
University of South Florida
Tampa, Florida, USA

Tuomo Tikkanen
Finnish Psychological Association
Helsinki, Finland

Bernhard Wilpert
Technische Universität Berlin
Berlin, Germany



Section Editors

Neil Anderson
Universiteit van Amsterdam
Amsterdam, The Netherlands

Andrew Baum
University of Pittsburgh
Pittsburgh, Pennsylvania, USA

John W. Berry
Queen's University
Kingston, Ontario, Canada

Mirilia Bonnes
Università degli Studi di Roma "La Sapienza"
Rome, Italy

Heliodoro Carpintero
Universidad Complutense de Madrid
Madrid, Spain

Giuseppe Carrus
Università degli Studi di Roma "La Sapienza"
Rome, Italy

Patrizia Catellani
Catholic University of Milan
Milan, Italy

Florence Denmark
Pace University
New York, New York, USA

Rocio Fernández-Ballesteros Garcia
Autónoma Universidad de Madrid
Madrid, Spain

Jane Goodman-Delahunty
University of New South Wales
Sydney, New South Wales, Australia

Yuri Hanin
Research Institute for Olympic Sports
Jyvaskyla, Finland

Ann Higgins-D'Alessandro
Fordham University
Bronx, New York, USA

Robert Hogan
Hogan Assessment Systems
Tulsa, Oklahoma, USA

Barry Kantowitz
University of Michigan
Ann Arbor, Michigan, USA

Uichol Kim
Chung-Ang University
Seoul, Korea

Howard M. Knoff
U. S. Department of Health and Human Services
Little Rock, Arkansas, USA

Frederick T. L. Leong
University of Tennessee
Knoxville, Tennessee, USA

Peter A. Lichtenberg
Wayne State University
Detroit, Michigan, USA

Ingrid Lunt
University of London
London, United Kingdom

Roy Malpass
University of Texas
El Paso, Texas, USA

Rick R. McCown
Duquesne University
Pittsburgh, Pennsylvania, USA

José M. Prieto
Universidad Complutense de Madrid
Madrid, Spain

Donald K. Routh
University of Miami
Miami, Florida, USA

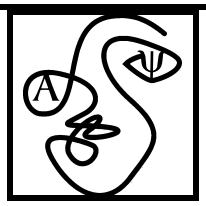
Mark L. Savickas
Northeastern Ohio Universities College of Medicine
Rootstown, Ohio, USA

Margaret Semrud-Clikeman
University of Texas
Austin, Texas, USA

Harry C. Triandis
University of Illinois, Urbana-Champaign
Champaign, Illinois, USA

W. Fred van Raaij
Universiteit van Tilburg
Tilburg, The Netherlands

Sheldon Zedeck
University of California, Berkeley
Berkeley, California, USA



Preface

During the past 50 years, there has been remarkable progress in the articulation of psychological theory, increasingly sophisticated reports of empirical research, and a dramatic expansion in professional applications of psychology in all areas of modern life. Applied psychology requires both a sound scientific foundation and attention to the practical problems that are encountered in applying psychology to an infinite variety of life settings. Growing public recognition of the potential contributions of psychology to solving life's problems has also fostered rapid growth in the number of psychologists who are engaged in research and in professional practice throughout the world.

The need for an authoritative publication that reviews and evaluates the numerous applications of professional psychology has been recognized for a number of years, and has been of special interest to the members of the International Association of Applied Psychology (IAAP), the world's oldest and largest membership organization of individual psychologists. The interests and concerns of IAAP members, who reside in more than 80 countries, have been strongly noted by former IAAP presidents Edwin A. Fleishman, Claude Levy-Leboyer, and Harry C. Triandis.

Significant contributions to international applications of psychology were explicitly recognized and discussed in the context of their diverse societal settings by Bernhard Wilpert in his Presidential address at the IAAP's 24th International Congress of Applied Psychology in 1998. It was my honor and pleasure to succeed Professor Wilpert as President of the IAAP at this Congress, which was held in San Francisco, California, and to begin participating in discussions to identify the most important contributions to applied psychology that are reported in this encyclopedia.

The development and current status of applied psychology can be best understood by examining the historical evolution of the field, which is cogently described by Helio Carpintero, the official Historian of the IAAP. In his overview article on the history of applied psychology, Carpintero highlights the benefits resulting from psychologists joining together to form professional associations that have facilitated the development and integration of psychological theory and research and that have stimulated and guided more effective applications of psychology in professional work.

Founded in 1920 as the *Association Internationale des Conferences de Psychotechnique* (International Association for Psychotechnology), the IAAP adopted its present name in 1955. Edouard Claprēde convened and chaired the first meeting of the association, held in Geneva, Switzerland in 1920, and served as its President for more than 20 years. Meetings were held in a number of European cities, including Barcelona, Milan, Paris, Moscow, and Prague. An excellent history of the evolution and development of applied psychology and the IAAP can be found in the address given by former President Edwin A. Fleishman on receiving the 1999 American Psychological Association's Award for Distinguished Contributions to the International Advancement of Psychology (*American Psychologist*, November, 1999, pp. 1008–1016).

For recommending that the IAAP take an active role in sponsoring and developing the *Encyclopedia of Applied Psychology*, we are indebted to George Zimmar, an active IAAP member and former senior sponsoring editor at Academic Press, now Elsevier. The first step in developing the encyclopedia was the appointment of an Advisory Board composed of distinguished colleagues representing all areas of

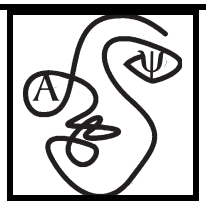
applied psychology. In addition to recommending the major areas and the specific topics to be covered in the encyclopedia, the Advisory Board identified highly qualified colleagues who would be invited to serve as Section Editors for the selected areas.

The responsibilities of the Section Editors included writing a major article to provide an overview and analysis of their respective areas and working with the Editor-in-Chief and the Advisory Board in identifying the specific topics to be included in each area. They also helped to determine the amount of space to be allocated for each topic and recommended colleagues with outstanding knowledge and expertise to serve as the authors for the individual articles. In addition, the Section Editors reviewed and edited the articles that presented information on the topics that were related to the subject matter that was included in their sections.

The articles that are included in the encyclopedia are presented in alphabetical order, as determined by the title of each article. Consequently, the underlying section organization will not be readily apparent. At the conclusion of each article, the reader is directed to additional articles that include related information. The authors of each article also provide a number of references that are recommended for further reading. It should be noted, however, that it was not possible to cover all of the many areas of applied psychology. Several emerging areas are not included because we failed to recognize their importance at the time the topics were selected for the encyclopedia. It is also unfortunate that a few authors were unable to meet the deadlines for submitting the manuscripts for their articles.

The *Encyclopedia of Applied Psychology* is truly a product of the collaborative efforts of the members of the IAAP. We are especially indebted to former IAAP Presidents Edwin A. Fleishman, Harry C. Triandis, and Bernhard Wilpert, and to our current president, Michael Frese, for their leadership, encouragement, and tangible contributions to the encyclopedia. For their dedicated contributions to the encyclopedia, I would also like to thank the members of the Advisory Board and the Section Editors, whose names are listed in the front matter. Finally, on behalf of the IAAP, I thank the authors who contributed articles in their areas of expertise and the editorial and production staff of Elsevier for the tremendous amount of time and energy that was needed to bring this large project to fruition.

CHARLES D. SPIELBERGER, Ph.D., ABPP
President, IAAP, 1998–2002



Personality and Emotion in Late Life

Nicole Alea, Manfred Diehl, and Susan Bluck

University of Florida, Gainesville, Florida, USA

1. Introduction
2. Personality in Late Life
3. Emotion in Late Life
4. Concluding Remarks: Linking Personality and Emotion
Further Reading

GLOSSARY

affect optimization Tendency to regulate emotional states toward more positive affective experiences and away from negative emotion.

cohort Group of people who are part of the same generation and, as a consequence, share similar life experiences and historical influences.

cohort-sequential design A study design that allows researchers to disentangle the effects of cohort and age as it tests different cohorts at the same ages across a period of time.

cross-sectional design A study design that allows researchers to draw conclusions about age differences by testing individuals of different ages at one point in time.

emotion regulation The ability to self-regulate or manage the internal experience and external expression of emotion.

emotions Short-lived negative or positive responses and reactions to events in the environment, or internal cognitions, that involve changes in body physiology, subjective experiences, and expressive behaviors.

life span development Development of individuals across the life span characterized by multidirectionality and multidimensionality; that is, developmental trajectories are assumed to include growth, stability, and decline within and across behavioral domains.

longitudinal design A study design that allows researchers to draw conclusions about age-related changes by examining the same group of individuals at different ages over a period of time.

personality Characteristics of an individual that are relatively stable and endure across time and settings.

primary emotions Set of emotions that are fundamental in organizing human thought and action across the life span, including happiness, fear, anger, sadness, surprise, disgust, and interest.

rank-order consistency The consistency of an individual's relative placement or rank within a group over time.

This article reviews the literature on personality development and emotional experience across the adult life span, with a particular emphasis on late life. It reviews theoretical developments and challenges, discusses methodological issues, and presents the major empirical findings about personality and emotion in late life. It concludes with a discussion of how future work that better integrates these two areas would be beneficial in advancing our understanding of well-being in late life.

1. INTRODUCTION

Advances in health care and public health practices have created a new culture of "old age." But what are old people like? Negative stereotypes suggest that older adults have certain personality characteristics, that is, that they are stubborn, set in their ways, and/or afraid

©2004 Elsevier Inc.
All rights reserved.

to take risks. Competing positive stereotypes suggest that older people are full of grace and dignity and are compassionate and wise. How do old people feel? Again, there are competing cultural conceptions. It is often believed that old people are lonely and depressed. On the other hand, old people are often viewed as carefree and satisfied. This article moves beyond cultural stereotypes of personality and emotion in later life to present the theories, methods, and findings gleaned from scholarly research.

Historically, many psychological characteristics (e.g., intelligence, learning ability) were characterized by decline in late life. More recently, however, psychologists have begun to focus on the positive experiences associated with aging. This dynamic tension between the positive and negative aspects of late-life functioning was elucidated in a 1987 article by Paul Baltes of the Max Planck Institute for Human Development in Berlin, Germany. He proposed the life-span approach and postulated that life span development is multidirectional and multidimensional. That is, different psychological characteristics show different developmental trajectories (e.g., growth, stability, decline) and different trajectories may coexist for behavioral functions within the same domain, making growth and positive development possible even into late life. Indeed, Smith's findings from the longitudinal Berlin Aging Study show that successful aging can be defined in terms of the ratio between gains and losses.

Personality and emotion have been at the forefront of research showing positive gains or maintenance of functioning in late life. As such, research on personality and emotion offers a window on facets of successful aging. This article reviews each area of work separately, focusing on major theories, methods, and empirical results.

2. PERSONALITY IN LATE LIFE

The term "personality" refers to the relatively stable characteristics of an individual that endure across time and settings. At first glance, this definition suggests that personality must be stable across adulthood; for example, an extraverted 30-year-old is expected to be extraverted when he or she is 65 years old. Evidence supports a good deal of continuity in adult personality, but there are also data showing that change occurs. Whether a researcher favors a view of stability or one of change depends mostly on his or her theoretical perspective.

2.1. Theoretical Issues

2.1.1. Personality Theories

Labouvie-Vief and Diehl, in a 1998 review of life span theories, pointed out that two major theoretical perspectives have led the area of personality development. The developmental-level approach is rooted in the psychodynamic tradition, with theorists such as Freud, Jung, and Erikson. This approach is categorized by developmental stages that emphasize qualitative changes in personality during different life periods. For example, Erikson theorized that each stage of development is characterized by a "psychosocial crisis" between opposing tendencies (e.g., generativity vs stagnation during midlife) that must be integrated for the self to fully develop. Although Freud's theory emphasized that personality was set at a fairly young age, later modifications (e.g., by Erikson and Loevinger) and extensions included lifelong processes of personality development.

The trait approach assumes that personality is composed of a number of basic behavioral tendencies (i.e., traits) and has focused on the measurement of these basic traits. McCrae and Costa's "Big Five" model of personality is one prominent example of this approach. According to this theory, there are five broad traits that make up human personality: Neuroticism, Extraversion, Openness to Experience, Agreeableness, and Conscientiousness. Trait theorists emphasize the stability of personality beginning in early adulthood.

A fairly recent addition to the personality literature is the contextual approach, which considers sociocultural and environmental influences that may affect how personality develops across the life span. Researchers such as Helson emphasize that people's environments and experiences influence who they become, at least in part. In contrast to the developmental-level and trait theories, the contextual approach argues for the malleability of personality across a lifetime. What are the mechanisms that lead to changes, or that sustain continuity, in personality?

2.1.2. Mechanisms of Continuity and Change

In a 1999 review, Caspi and Roberts summarized mechanisms that support continuity and those that facilitate change in personality across the adult life span. Genetics are thought to be a major mechanism leading to continuity. Data such as the MacArthur Longitudinal Twin Study suggest that approximately 80% of personality stability can be attributed to genetic

factors. Notable, however, is that genes seem to become less influential in late life. As twins age, they become less similar, suggesting that environmental influences play an increasingly stronger role as individuals age.

Environmental factors, such as differences in life experiences and socialization, are the mechanisms that are thought to produce changes in personality. For example, in a classic study, Helson and Moane showed that interpersonal dominance tends to increase in women who actively pursue a professional career, whereas raising children tends to heighten a woman's level of empathy. Thus, social roles that carry implicit and explicit demands can shape personality during adulthood. Individuals can also learn new behavioral patterns, at any age, by observing others or by getting feedback from others. This type of social learning refutes the old adage that "You can't teach an old dog new tricks." However, social and physical environments can also play a role in maintaining continuity. To the extent that people's environments remain stable, they foster stability in personality. Consistent environments tend to encourage consistency in personality across the life span. Thus, genetics and environment contribute to personality continuity and change in a complex integrated fashion. Researchers' theoretical perspectives, as well as their views of continuity and change, affect the questions they ask and the methods and measures they use.

2.2. Methods and Measures

Early theorists, such as Allport, Cattell, and Murray, suggested that personality should be studied over time. During more recent years, Block has advocated studying personality "the long way," that is, using longitudinal methods. Although the majority of personality research is still cross-sectional in nature, the number of longitudinal studies has risen considerably during recent years and invaluable information has been obtained regarding personality continuity and change in late life.

2.2.1. Types of Continuity in Longitudinal Research

Everyday conceptualizations of continuity and change in personality are typically thought of at the individual level, that is, involving intraindividual variability in personality (referred to as ipsative continuity). For example, has grandpa changed during the past 10 years, or is he the same person he used to be? Investigating personality in this way has only recently received adequate attention because doing so requires complicated longitudinal designs of intraindividual change and advanced statistical analyses. Rather, researchers have focused their attention on two other types of continuity. Differential continuity refers to the consistency of an individual's relative placement within a group over time, that is, rank-order consistency. For example, in the case of high rank-order consistency, individuals are expected to preserve their rank over time, whereas low rank-order consistency is indicative of the fact that people are shifting around in terms of their rank. The second commonly studied continuity is absolute continuity, which refers to stability in the level of a particular personality characteristic over time. It is important to note that when most individuals change in the same direction, changes in mean level can occur over time without great changes in rank-order consistency. Thus, there can be various types of continuity, or lack thereof, observed in longitudinal assessments. There are also various factors that affect whether change or continuity is observed.

2.2.2. Factors Affecting Whether Continuity or Change Is Observed

Recent meta-analyses conducted independently by Ardelt and by Roberts and DelVecchio identified similar factors affecting whether change or continuity in personality is observed in longitudinal studies. Three factors tend to bias results toward stability. The first factor is an individual's age at the initial time of testing. The older people are when personality is first assessed, the greater stability coefficients tend to be. Second, length of retest interval can affect findings. The closer together the times of testing, the more consistent personality appears. Third, the type of measurement instrument can also bias results. Self-report checklists of personality traits, the most commonly used instruments with older adults, tend to overestimate personality stability because individuals are usually biased in favor of reporting continuity. Behavioral observations get around the self-report bias in that they involve observations of how people behave across contexts. Given their contextual nature, behavioral observations are more sensitive to changes in personality, but they have rarely been used with older adults.

Other factors have an effect on whether change in personality is likely to be observed. Life events and transitions (e.g., retirement, widowhood) that individuals go through while participating in a longitudinal study may suggest change in personality. That is, as

individuals respond to new role demands, they may be learning new ways in which to cope. For example, financial strain that can accompany retirement may lead a person to become more worried (with "worrying" being part of the trait of Neuroticism) than they usually are due to the loss of income. This may be a new behavior that emerged as a result of the social transition, and it becomes incorporated into the individual's personality if reinforced over time or may be a temporary way of dealing with a new life transition. Unknowingly, researchers would observe changes in personality.

Historical experiences and generational effects can also affect whether personality change is observed. In his classic work about the children of the Great Depression, Elder demonstrated how historical events affect people's personalities in different ways depending on their life stages. Elder found personality change for middle-aged men in reaction to the Depression. Personality continuity was more often found for younger men because these men were able to seize newly emerging opportunities after the Depression. However, conclusions about personality development were confounded with individuals' birth cohorts; thus, it was unknown whether the findings were a result of maturational change, the time of testing and historical influences, or the participants' generations. Schaie developed the cohort-sequential design to tease apart the effects of these factors, yet this elaborate design is rarely used. In sum, longitudinal studies, although offering the most comprehensive picture of personality development, can be limited about personality continuity and change. Thus, when reviewing empirical results regarding personality across adulthood in the section that follows, one should pay close attention to the types of continuity found and the methods employed.

2.3. Major Findings

Table I summarizes the types of continuity and change evident in research on personality in late life, the typical statistical methods used, and the general conclusions. This section reviews major empirical results supporting both continuity and change in personality.

2.3.1. Evidence for Continuity

There is quite a bit of evidence pointing toward continuity of personality during adulthood and late life. Costa and McCrae's Baltimore Longitudinal Study began in 1980 and consists of approximately 1000 participants ages 20 to 96 years. Participants were given the NEO Personality Inventory, which measures the Big Five basic personality traits. In this and other work, they have found moderate to high rank-order consistency for personality traits across time intervals of 5 to 6 years. Costa and McCrae interpreted their results as indicating that personality does not change, except in trivial ways, after the age of 30.

Other researchers espousing the individual differences approach, such as Conley, have also found a great deal of stability in personality across adulthood. Using a different self-report personality measure as part of the Kelly Longitudinal Study, Conley and colleagues found that certain personality traits, such as neuroticism, social introversion–extraversion, and impulse control, demonstrated fairly stable patterns across 25- and even 40-year time intervals. Focusing on late life, Troll and Skaff found continuity in self-concept in the oldest old (85 years of age or over). Participants were asked how they thought they had changed over the years and how they had remained the same. These older adults reported that their core self-concept had remained stable over the years, and their self-descriptors were also relatively consistent. In sum, for different measures used to assess personality, there seems to be differential consistency of personality during adulthood and late life.

TABLE I
Summary of Personality in Late Life: Types of Continuity, Typical Methods, and Empirical Results

Types of continuity	*Typical methods*	*Empirical results*
Differential continuity	Rank-order consistency	Stability: ranks of personality characteristics relatively stable in late life
Absolute continuity	Mean levels	Change: decreases in neuroticism, increases in agreeableness, decreases and maintenance in extraversion in older adults
Ipsative continuity	Intraindividual differences and variability	Stability and change: growth, stability, and decreases in personality characteristics within individuals

2.3.2. Evidence for Change

Advocates of the contextual view, such as Helson, have argued that trait theorists seek to confirm hypotheses about the stability of personality rather than to explain why and how changes in personality occur when they do. In the Mills College Longitudinal Study, Helson found changes in adult women's personalities across a 30-year time period. For example, women became more assertive (i.e., mean-level increase on the masculinity scale) from 20 to 30 years of age, and they became more compassionate (i.e., mean-level increase on the femininity scale) as they approached late life (52 years of age). But what happens in very old age? Field and Millsap analyzed data from the Berkeley Older Generation Study and found consistent age-related trends for the mean level of some personality traits. For example, for this very old sample, they found increases in agreeableness and decreases in extraversion with age. Thus, there seem to be changes in mean levels of personality characteristics, even though rank-order consistency is fairly constant. Studies of intraindividual variability also suggest both change and continuity in personality across individuals' lives.

2.3.3. Intraindividual Variability: Evidence for Continuity and Change

Although few studies of intraindividual variability have been conducted, they provide a comprehensive picture of the variability of personality in late life. A classic study by Jones and Meredith examined individual patterns of personality development over a 30-year time span. For example, they found that some participants' self-confidence was stable until 30 years of age and then increased during middle age, others' self-confidence remained stable across the time period, and others' self-confidence declined with age. In a recent study, Mroczek and Spiro found variability in the rate and direction of change in personality among men ranging in age from 43 to 91 years. Although they found that many men showed declines in neuroticism and stability in extraversion, there were also many men who deviated from this overall pattern, thereby demonstrating individual variability even in old age.

In conclusion, personality in late life is undoubtedly characterized by both differential continuity and changes in mean levels of personality characteristics (i.e., absolute continuity). Studies of intraindividual variability reiterate that continuity and change in personality both occur in late life differentially for individuals and that group means can be deceiving. However, mean-level changes that occur are relatively positive (e.g., decreases in neuroticism, maintenance in extraversion). Multiple dimensions and directions of change characterize personality in late life. Does a similar variable pattern exist for how people feel in late life?

3. EMOTION IN LATE LIFE

Emotions have been defined as short-lived responses to events in the environment that involve changes in body physiology, subjective experiences, and expressive behaviors. That is, emotions are transient and happen on multiple levels. Some emotions, such as happiness, do not endure over a lifetime but rather are dynamically and subtlety tied to specific events and experiences—to cognitions, memories, and personalities. The important questions for emotion research have been whether older adults, given their unique life circumstances, cognitive abilities, and personalities, experience or feel emotions to the same degree as do younger adults, experience the same types of emotions as do younger adults, and regulate and express emotions in the same way as do younger adults. The following theories of emotion in late life address these central questions.

3.1. Theoretical Issues

3.1.1. Emotion Regulation and Affect Optimization

In 1989, Lawton proposed a theory of the regulation processes of emotion during adulthood, emphasizing positive gains in regulation with age. Furthermore, he postulated that in late life, adults regulate emotion toward "affect optimization," that is, a minimization of negative affect and a maximization of positive affect. He contended that better affect regulation in late life occurs as a function of positive changes in personality as well as changes in social contexts in late life. Examination of social context became a focal area for further theoretical expansion, as seen in the next theory.

3.1.2. Socioemotional Selectivity Theory

In 1992, Carstensen proposed the socioemotional selectivity theory, which situates emotion in the context of social goals associated with different life phases. This theory focuses on the salience of emotion in older adults' everyday lives. Carstensen, like Lawton, emphasized

regulating emotion toward obtaining positive affective experiences. The core of Carstensen's theory is that older adults show a selective narrowing of social networks to more emotionally close, gratifying social relationships. That is, the regulation of emotion is, in part, a function of social choices; maintaining a positively valenced emotional life is realized through selection of positive social relationships. Carstensen's theory ties emotion regulation to life phase-specific choices regarding social networks. The next theory also ties emotion to life phase, but largely through the relation of emotion to cognition and personality at different points during the life span.

3.1.3. Cognitive–Affective Developmental Theory

Labouvie-Vief proposed a cognitive–affective developmental theory of emotion that delineates the role of emotional functioning in conjunction with the development of cognitive and ego processes (the self) that occur across the life span. She argued that emotion is qualitatively different at various points in the life span. Specifically, with more complex ways of processing information and higher levels of ego development, emotional expression moves from being based on societal standards during adolescence to more sophisticated emotional functioning during adulthood. This includes increases in the complexity of experienced emotion, heightened emotional flexibility, and more tolerance for ambiguous emotions. Labouvie-Vief's theory clearly delineates the connection between emotion and personality, or self-development, as does the discrete emotions theory advocated by Magai.

3.1.4. Discrete Emotions Functionalist Theory

During the 1980s, Magai integrated emotion theory with personality theory by introducing the discrete emotions functionalist theory to research on adult development. Rooted in the differential emotions tradition, this theory emphasizes the functional nature of the primary emotions (e.g., happiness, fear, anger, sadness), focusing on how emotions are used to organize thought and behavior. Magai's work suggested that discrete emotions are hardwired from birth but serve similar functions across the life span. What changes with age is the ways in which emotions are expressed; that is, with increasing age, the complexity and elaboration of emotions increase. In addition, Magai delineated how emotion and personality are intertwined, particularly in late life. She suggested that discrete emotions experienced repeatedly early in life become traitlike and integrated as "emotion traits" within an individual's personality.

These four theories make unique contributions to the literature regarding how emotion is regulated toward positive affect in late life, how social processes are used to regulate emotion in late life, how life span development of self and cognition encourages more sophisticated processing of emotion with age, and how emotions play a functional role in directing thought and behavior from early childhood but gain increasing complexity and elaboration across adulthood. Despite the differences among these four theories, each theory in some way addresses one of the central issues of how older adults experience emotion, regulate emotion, and express emotion.

3.2. Methods and Measures

Corresponding with the central questions of emotion research, most work has been cross-sectional rather than longitudinal. Emotion researchers use a rich array of methodological tools, including both objective and subjective measures.

3.2.1. Objective Measures

Objective assessments of emotion are psychophysiological measures tied to the basic neurology of emotion (e.g., autonomic nervous system activity, discrete musculature changes during facial expression). Studies assessing the psychophysiological aspects of emotion involve asking participants to look at emotional pictures (e.g., a child playing in a field) or to relive emotional experiences (e.g., a surprising event). Autonomic activity, such as changes in heart rate or breathing, is assessed online during these tasks. Similar stimuli are used when examining facial expressions of emotions. Ekman developed an elaborate system for coding universal, emotion-specific facial expressions. For example, people universally express sadness by pulling down the corners of their lips, and surprise is expressed by a widening of the eyes. These objective measures of emotion inform researchers about the underlying processes and universal nature of emotions.

3.2.2. Subjective Measures

Subjective measures of emotion are more commonly employed. They typically include self-report ratings and, more recently, behavioral observations (e.g., coding the types of emotion present during social exchanges) and

open-ended verbal narratives (e.g., coding emotional content of remembered narratives). Self-report measures (e.g., the Positive and Negative Affect Scale developed by Watson and colleagues) are undoubtedly the most common means of assessing emotion in late life. Adults answer questions about whether, and how intensely, they have felt several negative and positive emotions (e.g., sadness, happiness) over the past month or the past week. Self-reports are subject to response biases and, as Levine and Bluck demonstrated, have a memory component that may lead to greater biasing in older adults. For example, questions about general emotional experience (e.g., "How salient is emotion to you in your everyday life?") probably require effortful processing if a person actually recalls everyday events to provide an estimate of the general emotional content. Thus, results using such scales are prone to memory error. More naturalistic and ecologically valid measures, such as observing emotional behaviors and collecting emotional memory narratives, may be better suited as subjective measures of emotion in late life. Bluck and Alea recently found that autobiographical narratives and scalar self-reports sometimes provide divergent information about the emotional lives of older adults.

In sum, both objective and subjective measures have been used to examine various aspects of emotion in late life. In terms of subjective measures, there are a number of tools that vary in level of reliability and ecological validity. The next section reviews major findings about emotional functioning during adulthood and late life, paying special attention to the type of measurement employed.

3.3. Major Findings

The major empirical findings mirror the chief theoretical concerns in this area: the overall experience of emotion, the regulation of emotion, and the expression of emotion during adulthood and late life. Table II summarizes the major empirical findings by method of investigation.

3.3.1. Experience of Emotion

Research on the experience of emotion includes both physiological measures and the subjective experience of emotion in daily life. Levenson and colleagues have conducted a series of studies investigating whether older adults experience the same intensity of physiological reactions as do younger adults when they relive emotional experiences. Using online physiological assessments, these researchers have found that, for the most part, older adults exhibit the same physiological responses as do younger adults for discrete emotions. For example, both younger and older adults have increases in heart rate when experiencing fear. The magnitude or intensity of these responses, however, is often lower in late life.

The intensity of experienced emotion has also been examined using self-report measures. For example, Diener and colleagues conducted a large-scale study to assess age differences in the self-reported intensity of emotions in adolescents, younger, middle-aged, and older adults. They used the Affect Intensity Measure, which assesses how intensely individuals experience emotions in various situations (e.g., "I can remain calm even on the most trying days"). Older adults reported less intense emotional experiences than did younger adults. Thus, older adults tend to show the same pattern of physiological reaction as do younger adults, but emotions generally seem to be experienced with less intensity in late life. This holds for both physiological measures and self-reports. But what

TABLE II
Summary of Emotion in Late Life: Aspect of Emotion, Typical Methods, and Empirical Results

Aspect of emotion	*Typical methods*	*Empirical results*
Emotional experience	Psychophysiological	Mixed: no age difference for physiological experience of primary emotions; lower intensity with age
	Self-reports	Lower intensity with age
Emotion regulation	Self-reports	Increase in positive affect and decrease in negative affect with age; better emotion regulation with age
Emotional expression	Facial expression	Mixed: no age differences; more expressive with age; harder to decode expressions with age
	Verbal expression	Express more emotion with age
	Behavioral indicators	Fewer negative behaviors and more positive behaviors

happens if one looks at emotional valence? That is, do older adults experience both positive and negative affect with less intensity than do younger adults?

3.3.2. Emotion Regulation

Empirical work investigating the balance between positive and negative affect with age has often been conducted in the context of emotion regulation. Most (but not all) work on emotion regulation is congruent with existing theoretical accounts claiming that older adults tend to regulate their emotions toward affect optimization. Older adults self-report experiencing more positive and less negative affect than do younger age groups. For example, Mroczek and Kolarz analyzed data from more than 2000 adults ranging in age from 25 to 74 years. Using the Bradburn Affect Balance Scale, a self-report scale of positive and negative affect, they found an increase in positive affect after 45 years of age and a linear decrease in negative affect across age groups. Charles and colleagues found similar results concerning negative affect in a 14-year longitudinal study of self-reported positive and negative emotion. Furthermore, Gross and colleagues have demonstrated that increases in positive affect states and decreases in negative affect states also apply to discrete emotions such as happiness and anger. Thus, although researchers have found overall decreases in the intensity of emotion with age, it seems to depend on whether the emotion is positive or negative. In late life, negative emotion is often down-regulated, whereas positive emotion is often up-regulated.

In addition to examining the regulation of emotion through valence, researchers have explicitly asked adults about their ability to self-regulate emotions. In the study by Gross and colleagues, emotion regulation was investigated in a diverse sample (e.g., European Americans, African Americans). Participants were asked how well they could control both their inner and outer expressions of emotion (e.g., "Overall, how much control would you say you have over your emotions?"). They found that across age groups, there was an increase in self-reported ability to regulate emotion. Using a different technique, Labouvie-Vief and colleagues reported similar results as far back as 1989. They asked participants to recall emotional experiences and then conducted semistructured interviews about the types of emotional control or regulation strategies individuals used. They found age differences in emotion regulation, although the primary difference was between adolescents and the two older age groups (middle-aged and older adults). Thus, using self-reports, there is reasonable evidence suggesting that older individuals are better at regulating emotions than are younger individuals and that older individuals tend to regulate their emotions toward positive affective states. Emotion regulation usually refers to internal experiences of emotion. Does emotional expression differ by age?

3.3.3. Expression of Emotion

Emotion can be expressed in three basic ways: facial expressions, verbal expressions, and specific behaviors. Levenson and colleagues examined the facial expressions of young and older adults while they were remembering and describing emotional experiences. The researchers failed to find age differences in how adults expressed the primary emotions of anger, disgust, fear, happiness, and surprise. In contrast, Magai found that older adults were at times more expressive than younger adults but that physical changes in their faces with age (e.g., wrinkles) can mask their expressivity.

Research examining the verbal expression of emotion has also found that older adults express greater emotion than do younger adults. In 1994, Carstensen and Turk-Charles conducted a study in which they asked younger and older adults to remember narrative passages. Compared with younger adults, older adults spontaneously focused more on the emotional aspects of the narratives than on the factual information. Alea and colleagues have found similar age results in their work using autobiographical narratives. Specifically, older adults remembered more emotional information about an autobiographical event than did younger adults.

Behavioral observations of emotional expression, common in the child development literature, have been less common in studies of adulthood. Emotionally expressive behaviors include gestures (e.g., banging one's fists when angry, crying when sad) and body movements (e.g., looking away from something in disgust, shifting when nervous). Carstensen and colleagues conducted one of the few studies examining such behaviors in older adults. These researchers videotaped older and middle-aged married couples discussing positive, negative, and neutral marital issues. They found that, after controlling for level of marital satisfaction, older couples showed less negative emotional behaviors during discussions of unpleasant topics and were more affectionate toward their partners than were middle-aged couples. Thus, work on emotional

expressivity suggests greater expressivity in older adults' verbalizations and behaviors, although the basic facial configurations used to express emotion do not appear to change with age.

In sum, empirical results regarding emotional functioning in late life are centered on the issues of experience, regulation, and expression of emotion with age. Using various measures has provided a rather complex yet comprehensive picture of emotion in late life. In general, emotions sometimes are experienced less intensely with age and at other times are experienced just as intensely in late life as in earlier life; however, for the most part, emotions are regulated toward positive affect and expressed in a variety of ways across the life span. Thus, the pattern of emotional functioning during adulthood is also mostly characterized by gains and maintenance of functioning in late life.

4. CONCLUDING REMARKS: LINKING PERSONALITY AND EMOTION

This article began with a discussion of the competing negative and positive cultural stereotypes about older adults' personalities and emotional lives. Does scholarly work support or refute these different pictures of late life? Theoretical and empirical work on personality development and emotional functioning suggests that although older adults certainly face losses (e.g., widowhood) and challenges (e.g., increases in physical ailments), they also have psychological resources available to compensate for, or moderate, losses in functioning. Personality is mostly stable in late life, and (on the whole) changes seem to be relatively positive. Changes in emotional functioning with age, such as better emotion regulation and decreases in the intensity of physiological arousal levels, may also help individuals to face challenges in late life. Thus, both personality development and emotional functioning represent areas of stability or growth that enhance the experience of growing old.

Consequently, a clear step toward better understanding well-being in late life is further delineating the integral relation between personality and emotion. Although the major personality theories imply that traits have an inherent emotional component, the assessment of personality and emotion linkages has been rather sketchy so far. Emotion theories have been better at delineating the relation between personality and emotion. For example, Labouvie-Vief's work suggests that emotional development occurs during adulthood in part due to development of the self, and Magai's work postulates that emotions that are repeatedly experienced early in life become salient to an individual and integrated into his or her self-representations across the life span. Recently, with the publication of Block's book *Personality as an Affect-Processing System,* there is a growing awareness that the traditional distinct notions of personality and emotion may limit our understanding. Block suggested that these approaches can be fused in a creative way to examine both phenomena more comprehensively while including a life span developmental perspective relevant to the study of aging. The task for future researchers is to explore empirically how emotion and personality are interrelated and how they act together to influence quality of life in late life.

See Also the Following Articles

Aging and Competency ■ Anxiety Disorders in Late Life ■ Cognitive Aging ■ Depression in Late Life ■ Elder Caregiving ■ Emotion ■ End of Life Issues ■ Psychotherapy in Older Adults

Further Reading

Baltes, P. (1987). Theoretical propositions of life-span developmental psychology: On the dynamics between growth and decline. *Developmental Psychology*, 23, 611–626.

Block, J. (2002). *Personality as an affect-processing system: Toward an integrative theory*. Mahwah, NJ: Lawrence Erlbaum.

Carstensen, L. L. (1992). Motivation for social contact across the life span: A theory of socioemotional selectivity. In J. E. Jacobs (Ed.), *Nebraska Symposium on Motivation* (pp. 209–254). Lincoln: University of Nebraska Press.

Caspi, A., & Roberts, B. W. (1999). Personality continuity and change across the life course. In L. A. Pervin, & O. P. John (Eds.), *Handbook of personality: Theory and research* (2nd ed., pp. 300–326). New York: Guilford.

Helson, R., & Kwan, V. S. Y. (2000). Personality development in adulthood: The broad picture and process in one longitudinal sample. In S. E. Hampson (Ed.), *Advances in personality psychology* (Vol. 1, pp. 77–106). Hove, UK: Psychology Press.

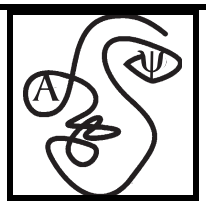
Labouvie-Vief, G., & Diehl, M. (1998). Life-span developmental theories. In C. E. Walker (Ed.), *Comprehensive clinical psychology*, Vol. 1: *Foundations* (pp. 261–296). Oxford, UK: Pergamon.

Lawton, M. P. (1989). Environmental proactivity and affect in older people. In S. Spacapan, & S. Oskamp (Eds.), *Social psychology of aging* (pp. 135–164). Newbury Park, CA: Sage.

Magai, C. (2001). Emotions over the life span. In J. E. Birren, & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (5th ed., pp. 399–426). San Diego: Academic Press.

McCrae, R. R., & Costa, P. T. (2003). *Personality in adulthood: A five-factor theory perspective* (2nd ed.). New York: Guilford.

Pasupathi, M., Carstensen, L. L., Turk-Charles, S., & Tsai, J. (1998). Emotion and aging. In H. S. Friedman (Ed.), *Encyclopedia of mental health* (Vol. 2, pp. 91–102). San Diego: Academic Press.



Personality Assessment

Gian Vittorio Caprara
University of Rome La Sapienza, Rome, Italy

Daniel Cervone
University of Illinois, Chicago, Illinois, USA

1. Modern Approaches and Contemporary Trends
2. Issues and Strategies
Further Reading

GLOSSARY

Big Five A set of personality dispositions (extroversion or energy, agreeableness or friendliness, conscientiousness, emotional stability or neuroticism, and openness to experience) considered by some to be present, to varying degrees, in every individual.

personality assessment Procedures designed to identify and evaluate enduring psychological qualities, including modes of thinking, feeling, and acting that characterize the individual and differentiate individuals from one another.

psychodynamic assessment Procedures, involving interviews and projective techniques, concerned with assessing the complexity of individual experience, including unconscious motives and conflicts among subsystems of personality.

Rorschach inkblot tests Cards with blots of ink (black, black/red, or multicolored) shown to subjects for interpretation. The subjects' interpretations of the abstract designs are considered projections indicative of their subconscious minds.

temperament Personality features evident early in life.

Thematic Apperception Test (TAT) Similar to a Rorschach inkblot test except that it is produced in shades of black and white and is somewhat more concrete, with an emphasis on psychoanalysis themes. Like the Rorschach test, the subjects' interpretations of the scenes are considered projections from their subconscious.

Personality assessment refers to procedures designed to identify and evaluate enduring psychological qualities, including modes of thinking, feeling, and acting that characterize the individual and differentiate individuals from one another. Assessment techniques address the cognitive, affective, motivational, and moral components of individual functioning, including cognitive abilities and styles, temperament and mood, motives, attitudes and values, habitual behaviors, coping strategies, and self concepts and regulatory mechanisms. Through a great variety of quantitative and qualitative techniques, personality assessors serve both basic and applied research needs in settings in which descriptive, predictive, and explanatory information about people is required. Personality assessments may be conducted for a variety of purposes. One may aim to describe normal variations in any of the previously mentioned individual characteristics in the population at large, to diagnose psychopathology suffered by a small subset of the population, or to obtain a detailed portrait of the psychological structure and dynamics that characterize a particular individual. The existence of widely varying techniques also reflects the presence in the professional field of competing theoretical views that one may endorse when pursuing an assessment of personality.

©2004 Elsevier Inc.
All rights reserved.

1. MODERN APPROACHES AND CONTEMPORARY TRENDS

Modern approaches to personality assessment can be traced to late 19th- and early 20th-century efforts to assess intelligence, temperament, and vulnerability to mental illness. While Francis Galton was an advocate of measurement of both intellectual abilities and character, Alfred Binet set the premises for the long-lasting Intelligence Quotient tradition, and Emile Kraepelin and Carl Gustav Jung pointed to word association as a common technique to detect and analyze mental disorders and complexes. The Woodworth Personal Data Sheet, which appeared in 1918, represents the earliest model of a paper-and-pencil self-report inventory to assess personality, whereas the Rorschach inkblot test, which appeared 2 years later, remains the most famous exemplar of projective techniques, which were designed to circumvent the psychological defenses that, in principle, might invalidate responses to self-report paper-and-pencil test items.

Two main lines of inquiry set the stage for subsequent developments in personality assessment. One primarily focused on interindividual genetic differences, whereas the other was mostly focused on the internal dynamics of individuality. Investigators who assessed phenotypic individual differences, namely traits, motives, and attitudes, conceived personality as an architecture of tendencies and habits. In contrast, theorists focusing on internal personality dynamics perceived personality as a system in service of adaptation and pointed to internal forces, external pressures, and defense mechanism that underlie individual personality development and shape its distinctive features of functioning.

During the subsequent decades, self-report trait inventories were commonly used in a variety of contexts to describe and predict performance and adaptation. Projective techniques such as the Rorschach inkblot tests and the Thematic Apperception Test (TAT) were more frequently used in clinical settings. Both were largely endorsed by scientists and practitioners, regardless of their theoretical orientation. Although there was not yet a unique science of personality, most progress came from work in psychometrics, including the extension of factor analytic methods to study the components of personality.

Thus, by the mid-20th century, personality assessment could count on not only a large variety of instruments but also a consistent body of knowledge regarding sources of data and the quality of measures. It was made clear that one should rely on different sources of data, such as behavioral observations, self-reports, psychophysiological indices, and narratives, to tap aspects of personality, and that alternative instruments may be differentially appropriate to capture general versus unique features of the individual, to assess different constructs, or to apply to different populations or in different settings. In this regard, the seminal contributions of Cattell to psychometrics and statistics and of Cronbach and Meehl to the quality of personality measures and to the various methods to ascertain their validity and reliability were long-lasting and influential.

However, sophistication of methods could not solve all disputes at the theoretical level. Today, although personality psychology is a well-established domain of scientific inquiry, the desirable idea of a unique science of personality in which multiple contributions from social, cognitive, psychodynamic, and dispositional orientations converge synthetically in the service of cumulative knowledge is far from being fully achieved. Instead, one finds a number of distinct broad trends in theory and assessment, each of which can claim to have made scientific progress.

One such trend addresses the subject as an active interpreter of his or her world. The cognitive and affective revolutions ultimately led to better appreciation of the role of the self system in granting unity, continuity, and coherence to individual experience. A new conception of personality as both an agent capable of making things happen according to its anticipations and standards and a social construct embedded in a network of interpersonal relationships has gained broad consensus. This widened the traditional view of personality as a mediated system, further highlighted the importance of continuous individual—environment interactions, reoriented the study of individual differences from global dispositions to contextualized behavioral strategies, and pointed to the systems that mostly account for the attribution of meanings and regulation of action as they operate and develop over the course of life.

Another area of advancement involves the assessment of personality features evident early in life. Research on temperament has begun to assess specific biological systems involved in the generation and regulation of affective experience. Much of this work has demonstrated that the assessment of temperament must consider not only the biology of personality but also the social contexts in which the person acts, especially because the nature and meaning of biological response measures may shift from one context to another.

Studies on emotional, social, and practical intelligence have brought closer the traditions of research on intelligence and personality that had often developed along separate pathways, the former focusing on cognitive abilities and problem solving and the latter on interpersonal relations and social behavior. The discovery of the cognitive unconscious, namely the variety of phenomena associated to automatic and implicit forms of knowledge, gave new impulse to the study of informational and motivational processes that operate beyond individual awareness.

The study of self systems has led to a focus on self-representations, self-beliefs, and self-regulatory mechanisms as they operate in concert to grant unity and coherence to individual experience and behavior. These changes in personality theorizing inevitably have carried new assessment aims, such as the assessment of skills and knowledge that underlie overt behavioral competencies, and have given impulse to the design of new procedures and instruments, including cognitive techniques to measure individual differences in knowledge accessibility and biopsychological techniques for tapping brain systems in personality functioning.

This progress has greatly benefited from major advances in the neurosciences that have yielded enhanced knowledge of brain functioning, as well as from the technological devices that allow closer access to the biological mechanisms underlying thought, affect, and action. In addition, access to large and diverse populations and computer-aided psychometrics further widened the scope of personality assessment process.

2. ISSUES AND STRATEGIES

Throughout much of its history a great part of personality psychology has been concerned with individual differences in observable variations in styles of behavior, affect, and cognition. Self-report and observer ratings have been the main methods to assess habitual behaviors, attitudes and beliefs, motives, and values.

A number of authors have posited traits to account for stable patterns of experience and action that people exhibit and that distinguish them one from another across tasks and situations. Within this line of thinking, personality has been conceptualized as a hierarchical organization with high-level traits (e.g., extroversion) that organize lower level tendencies (e.g., sociability), which in turn supervise lower level behavioral habits (e.g., talkative). Within this tradition, a broad consensus has been achieved on a central question, namely, the number and content of the high-level traits. Five global dispositions are seen as necessary to individual-difference assessment: extroversion or energy, agreeableness or friendliness, conscientiousness, emotional stability or neuroticism, and openness to experience. These or similar individual-difference dimensions have been consistently identified across different cultural contexts. These dimensions have proven to subsume many previous taxonomies, with a variety of specific traits being shown to correspond to facets or variations of each main trait or combination of main traits. Among the most commonly employed instruments to assess the so-called Big Five traits are the Revised NEO Personality Inventory of Costa and McCrae and the Big Five Questionnaire of Caprara *et al*.

During the past 20 years, many studies have corroborated the concurrent and predictive validity of the Big Five in a variety of educational, organizational, and health settings. Yet this approach meets only part of what personality assessment should aim to achieve.

The Big Five provide a sufficiently comprehensive description of personality at the level of surface behavior tendencies. They may serve to tap interindividual differences by pointing to features that may significantly impact on others as well as predict performance in a variety of settings. However, it is beyond the reach of the Big Five to explain the underlying mechanisms that govern behavior and grant its continuity and coherence. The constructs and their associated assessments indicate little about essential characteristics of being a person, such as self-awareness and intentionality. They meet the need of the bystander psychologist, but they miss the perspective of self-reflective agent.

In contrast, psychodynamic theories have long been much more concerned with assessing the complexity of individual experience. In this tradition, interviews and projective techniques such as TAT or the Rorschach continue to be used by clinicians as well as by researchers. Unfortunately, work in this tradition has encountered numerous empirical obstacles since these instruments are hardly amenable to standard criteria. Reviews indicate that in many important applications, projective tests fail to display adequate predictive validity.

Other investigators have sought alternatives to trait and psychodynamic assessment approaches that are able to capitalize on both quantitative and qualitative methods and to meet the requirements of objective research seeking for general laws without sacrificing what is unique of individual experience. One

theoretical grounding for such efforts has been social cognitive theories. In conceiving of personality as an open, dynamic, unifying, and integrating system, social cognitive theorists point to the emerging properties of mind and focus on the processes and mechanisms conducive to knowledge and motivational structures that enable personality to function as a proactive self-regulatory system. Persons and sociocultural environments are viewed as reciprocally interacting systems, with the core person variables being cognitive and affective mechanisms through which people interpret the world and regulate their own behavior. Thus, social cognitive theorists focus on the self-construction of personality as an integrative and coherent system as it takes place over the course of life, on the processes that enable the system to function proactively in continuous and reciprocal interaction with the environment, and on the self-structures that orchestrate these processes. Where assessment issues come to the fore, the primary targets are self-regulatory structures, including self-representations, self-efficacy beliefs, and personal goals and standards, namely the structures that are involved most centrally in the self-reflective and self-evaluative properties of mind that operate at the individual level to meet environmental opportunities and constraints and to maintain personally valued courses of action.

See Also the Following Articles

Extroversion–Introversion ■ Psychological Assessment, Standards and Guidelines for ■ Psychometric Tests ■ Traits

Further Reading

Aiken, L. R. (1996). *Personality assessment methods and practices*. Kirkland, WA: Hogrefe & Huber.

Bandura, A. (1986). *Social foundations of thought and action*. Englewood Cliffs, NJ: Prentice Hall.

Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology*, *52*, 1–26.

Caprara, G. V., Barbaranelli, C., Borgogni, L., & Perugini, M. (1993). The Big Five Questionnaire: A new questionnaire for the measurement of the five factor model. *Personality and Individual Differences*, *15*, 281–288.

Caprara, G. V., & Cervone, D. (2000). *Personality: Determinants, dynamics, and potentials*. New York: Cambridge University Press.

Cattell, R. B. (1946). *Description and measurement of personality*. New York: World Books.

Costa, P. T., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual*. Odessa, FL: Psychological Assessment Resources.

Cronbach, L. J. (1970). *Essentials of psychological testing*. New York: Harper.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*, 281–302.

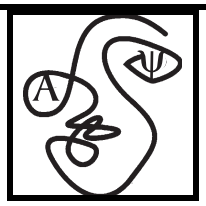
Fazio, R. H., & Olson, M. A. (2003). Implicit measures in social cognition research: Their meaning and use. *Annual Review of Psychology*, *54*, 297–327.

Kagan, J. (2003). Biology, context, and developmental inquiry. *Annual Review of Psychology*, *54*, 1–23.

Lilienfeld, S. O., Wood, J. M., & Garb, H. N. (2000). The scientific status of projective techniques. *Psychological Science in the Public Interest*, *1*(2).

Mischel, W. (1968). *Personality and assessment*. New York: Wiley.

Wiggins, J. S. (2003). *Paradigms of personality assessment*. New York: Guilford.



Personal Space

Robert Sommer
University of California, Davis, California, USA

1. Personal Space Concept
2. Measurement of Personal Space
3. Research Findings
4. Applications
Further Reading

GLOSSARY

individual distance The amount of space between organisms and their conspecifics.
personal space The emotionally tinged zone around the human body that people feel is "their space."
proxemics A term introduced by Edward Hall for the study of spatial relationships; Hall identified four interaction zones: intimate, personal, social, and public distance.
territory A fixed geographical space marked and defended by an organism and used for life-sustaining activities.

Because the concept of personal space originated in ethology, some of the most promising theories to explain why humans maintain predictable distances from one another employ a Darwinian (functional) framework. Researchers have investigated differences in interpersonal spacing associated with culture, gender, age, personality factors, and situation. Applications of the concept are found in design, communications, and legal cases.

1. PERSONAL SPACE CONCEPT

The term "personal space" was introduced into the social psychological literature to describe the emotionally tinged zone around the human body that people feel is "their space." It has been described using analogies to a soap bubble, a snail shell, and an aura around the human body. The dimensions are not fixed but rather vary according to internal state, age, culture, and context. The concept has its roots in animal studies, particularly the work of ethologists and zoologists, because most animals in the wild maintain defined distances from conspecifics, and these distances are influenced by an animal's age, size, gender, social rank, and other factors.

2. MEASUREMENT OF PERSONAL SPACE

Methods for measuring personal space in human interactions include both field studies and laboratory simulations (Figs. 1 and 2). A favorite laboratory technique is the invasion, which produces tension-reducing responses such as gaze aversion and withdrawal. In the stop distance technique, a researcher approaches the participant, who tells the researcher to stop when he or she comes uncomfortably close. In the approach distance technique, the participant is asked to move toward another person and to stop at a comfortable interaction distance. In figure placement tests, the participant places manikins, dolls, or other human

©2004 Elsevier Inc.
All rights reserved.



FIGURE 1 Field observation of seating in a college library.



FIGURE 2 Laboratory simulation of a space station environment.

surrogates in various social arrangements such as having a conversation. Other measurement techniques include questionnaires and physiological recording.

3. RESEARCH FINDINGS

Interpersonal distance is decreased by similarity between participants, acquaintanceship, attraction, cooperation, and cohesion. It is increased by threat, anxiety, stigma, smoking, and mental disorder. Children have smaller interaction distances than do adults, female pairs generally sit and stand closer than do male pairs, and people from Latin cultures use smaller interaction distances than do people from Anglo-Saxon backgrounds. Environmental variables such as noise, crowding, and room size affect interaction distance.

4. APPLICATIONS

There is a flourishing research literature on interpersonal spacing in human services fields such as social work, nursing, counseling, and family therapy. Guidebooks and self-help books describe appropriate spacing in various interpersonal encounters. Some books dealing with sales, management, and dating specifically recommend the aggressive appropriation of space (i.e., standing excessively close) as a form of impression management.

As personal space enters the popular culture, airlines advertise more of it in their seating, residents of homeless shelters complain that they have too little of it, and corporate training manuals warn employees to respect each other's personal space to protect the company against lawsuits.

4.1. Design and Architecture

The personal space concept has been applied in the design of offices, stores, banks, and other building types, but its greatest applicability is in mass transit and institutional settings with fixed seating and little opportunity for personal mobility. The U.S. space agency, the National Aeronautical and Space Administration (NASA), used the results of personal space research to improve habitability in the space station.

4.2. Communication Skills

Diplomats and corporate executives being posted to different cultures are taught about the different ways in which people around the globe use space and time. These ideas have also been applied to interactions between salespeople and customers, police interrogations, nurse–patient and therapist–client relationships, and interactions among family members.

4.3. Legal Uses

In the United States, the concept has had nearly as much application in the courtroom as in design. It has figured prominently in lawsuits on prison crowding and became a pivotal issue in cases of sexual harassment as unwanted closeness was interpreted as a form of harassment toward individuals who lacked power. Judges have used the concept in establishing "no intrusion zones"

between protesters and their adversaries. In high-profile trials, space consultants advise lawyers during jury selection, a time when potential jurors are being screened by both the prosecution and the defense for signs of bias. Nonverbal behaviors such as posture, gesture, and interpersonal distance are observed as potential jurors respond to lawyers' questions.

See Also the Following Articles

Environmental Stress ■ Privacy ■ Residential Satisfaction and Perceived Urban Quality

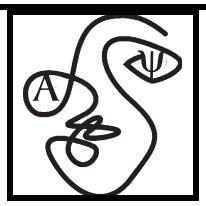
Further Reading

Aiello, J. R. (1987). Human spatial behavior. In D. Stokols, & I. Altman (Eds.), *Handbook of environmental psychology* (Vol. 1, pp. 505–531). New York: John Wiley.

Bechtel, R. B. (1997). *Environment and behavior*. Thousand Oaks, CA: Sage.

Sommer, R. (1969). *Personal space*. Englewood Cliffs, NJ: Prentice Hall.

Sommer, R. (2002). Personal space in a digital age. In R. B. Bechtel, & A. Churchman (Eds.), *Handbook of environmental psychology* (pp. 647–660). New York: John Wiley.



Person–Environment Fit

Cheri Ostroff and Kerstin A. Aumann

Columbia University, New York, New York, USA

1. Overview and Definitions
2. Creation of and Importance of Fit
3. Assessing and Measuring Fit
4. Selected Applications of Fit
5. A Potential Paradox of Fit
6. Conclusion
   Further Reading

GLOSSARY

attraction–similarity–attrition (ASA) cycle Cycle whereby individuals are attracted to, selected by, and remain in organizational environments that have characteristics similar to their own.

complementary fit The extent to which an individual's attributes provide something that is needed or lacking in the environment (e.g., the group, the organization).

levels of analysis Delineation of organizations into interrelated levels of analysis—the individual, the group or unit, and the organization as a whole.

occupational fit The extent to which an individual's vocational interests and skills match the opportunities and requirements of the occupation.

person–environment fit The extent to which the characteristics of an individual (e.g., skills, abilities, goals, needs, personality, values) match the characteristics of the work environment (e.g., job requirements, work group climate, organizational goals, organizational culture).

person–group fit The extent to which an individual's attributes (e.g., skills, goals, values, personality) are similar to or complement those of others in the work group.

person–job fit The extent to which an individual's skills and abilities match the skill and ability requirements of the job.

person–organization fit The extent to which an individual's attributes (e.g., goals, values, personality) are consistent with those of the culture or climate of the organization.

supplementary fit The extent to which an individual's attributes are congruent with or similar to those of the group or organization.

Person–environment fit theory addresses the extent to which the characteristics of an individual are similar to, or are complementary to, the characteristics of the work context, group, or organization. The degree of "fit" has implications for individual attitudes and behaviors in organizations, as well as for group and organizational effectiveness. Fit can be created through naturally occurring processes such as a cycle of being attracted to, selected by, and remaining in an organization, or through deliberate interventions such as socialization and training. Fit has a broad range of applications in organizations, including culture and climate, selection and socialization, and team composition and diversity.

1. OVERVIEW AND DEFINITIONS

Consider the hiring practices of two different organizations in the same manufacturing industry. In the first organization, the selection process entailed the

©2004 Elsevier Inc.
All rights reserved.

screening of each applicant's résumé to assess relevant background and work experience, a written test to assess job knowledge, and a short interview. This type of hiring process is not unusual in many manufacturing organizations. Now consider the hiring process used at Toyota when it opened a new plant in Kentucky. The process entailed a general knowledge test, a manufacturing exercise to provide a realistic job preview of the type of work to be performed, surveys of attitudes toward work, extensive interviews, and a physical exam. Each applicant spent at least 18 hours engaged in the selection process. Why would an organization devote so much time, money, and resources to hire employees from the lowest to highest job levels?

In the first case, the organization's primary concern was to achieve person–job (P–J) fit, that is, to hire employees with the relevant skills and abilities for specific jobs in the organization without specific consideration of the extent to which job applicants' characteristics are congruent or compatible with the characteristics of the organization in which the job is embedded. In the second case, the intensive selection process was geared toward hiring "whole" persons, that is, those who not only would fit the job but also would fit well into the organization's culture and the organization as a whole. Organizations that engage in such intensive hiring practices maintain that P–J fit alone is not sufficient to ensure that an employee will perform well in the organization. Rather, they advocate that employees who fit the job and the environment of their work group and the organization overall will be better performers, more satisfied employees, and less likely to leave the organization.

1.1. Definition: What Is Fit?

The fit or match between the characteristics of the individual, on the one hand, and the characteristics of the environment, on the other, is what is generally referred to as person–environment (P–E) fit. On the person (P) side, these characteristics may include skills, abilities, personality, personal needs, goals, and interests. On the environment (E) side, these characteristics may include specific job requirements, reward systems, goals, culture, climate, and resources of the unit or organization.

The underlying assumption is that a good fit between the characteristics of people and those of the environment will result in better performance, higher job satisfaction, enhanced loyalty to the organization, more positive attitudes, and a lesser desire to leave the organization for another job. Furthermore, it is assumed that fit is a dynamic and developmental process such that a lack of fit will prompt a change in careers, jobs, groups, or organizations in the hope of achieving fit. Although this notion may appear to be intuitive, understanding how fit operates in organizational settings is complex. For example, what processes help to create fit? How can an organization best identify and attract those individuals who will better fit the organization? How can an organization assess which job applicants will ultimately be a good fit for the job, the work group, and the values and culture of the organization? Is fit always desirable? These types of questions represent the core of fit research and theory and are addressed in the following sections.

1.2. Brief History: From Interactional Psychology to Fit

Researchers in organizational behavior are concerned with understanding and predicting people's responses in organizational settings. Historically, researchers debated whether individuals' responses are primarily a function of their personal attributes (e.g., personalities, knowledge, past experiences) or the characteristics of the situation (e.g., organizational climate, rewards system, organizational structure). Interactional psychology theories assume that individuals and their situations are interdependent. The characteristics of the situations constrain and affect the responses of individuals, whereas individual responses simultaneously influence characteristics of their situations or organizations. In other words, people react to situations and also create or enact them.

Although the interactionist perspective emphasizes the importance of both personal and situational characteristics, little theoretical explanation as to how person and environment variables interact to affect behavior was provided. The principle of congruence or fit is one explanation that has gained considerable recognition, and it highlights that person and environment factors should combine through "matching." The optimal combination is one in which the personal and situational characteristics are matched or compatible. Conditions failing to reach this kind of match will lead to a state of incongruence that will have negative effects on outcomes such as performance and satisfaction.

This notion that a better fit between people and their environment results in more positive behaviors and responses has a long history in various fields of psychology, particularly social psychology, personality, group

dynamics, and educational psychology. For example, in social and personality psychology, laboratory experiments over the past 40 years or more have suggested that for each individual, there are environments that more or less match the characteristics of the individual's personality. In educational psychology, seminal studies beginning in the 1950s and 1960s demonstrated that larger discrepancies between self-ratings and ideal ratings of the college environment and between self-ratings and other students' ratings of the environment were related to greater dissatisfaction with college life. In vocational psychology, a substantial amount of research has been devoted to Holland's seminal theory of vocational choice, indicating that greater career success and satisfaction with work purportedly results from a match between individuals' interests and the tasks and activities associated with broad classes of occupations (as opposed to specific attributes of a particular job or a particular organization).

The early work in industrial and organizational psychology research domains addressed (implicitly if not explicitly) the notion of fit to job characteristics. In particular, a great deal of emphasis was placed on personnel selection, that is, devising selection procedures and tests to maximize the probability of an organization hiring individuals whose skills and abilities matched the skills and abilities required for the particular job. Through training, attempts were made to maximize fit by upgrading the skills of individuals to better match the job requirements. Furthermore, in the area of job design, emphasis was placed on redesigning jobs to fit the abilities and limitations of the people performing them. This early work focused largely on fit as it pertained to skills and abilities or the specific type of attributes and tasks inherent in the job itself. Surprisingly little attention was directed at understanding fit in the broader social, work group, and/or organizational contexts.

Starting in the 1980s, theory and research on fit in organizational psychology domains became more complex and integrative. There was a shift from a predominantly skill-oriented P–J fit perspective to a broader perspective of addressing fit between a wide range of individual attributes and a wide range of job, work group, and organizational characteristics across a wide range of organizational areas.

1.3. Fit to Different Levels of Analysis

The primary emphasis in organizational fit research has been on the extent to which individuals fit the attributes of their job, their group, or their organization. As a result, three broad types of fit have been proposed. P–J fit focuses on congruence or fit between an individual's skills and abilities and the skills and abilities required to perform a specific job in the organization. Person–group (P–G) fit addresses the extent to which an individual's characteristics (e.g., personality, skills, values, goals) are congruent or aligned with the characteristics of other members of his or her work group or work unit. Fit to the broader organizational context is addressed through person–organization (P–O) fit focusing on the extent to which an individual's personality, goals, values, and attitudes fit in with the organization's broader context characteristics such as organizational climate, goals, values, policies, and systems.

Theory suggests, and some research has begun to illustrate, that aggregating across multiple types of fit (e.g., P–J and P–G together) is related to more positive responses (e.g., job satisfaction, reduced stress) than when fit is assessed by only a single type of congruence. On the other hand, some have argued that certain types of fit may be more important than others. For example, when evaluating the extent to which they fit, individuals tend to place a heavier emphasis on fit with salient environmental aspects than on fit with less salient environmental aspects. To date, little research has explored the relative importance of different types of fit.

1.4. Supplementary and Complementary Fit

The preceding conceptualizations define fit broadly as the congruence between individuals' characteristics and those of the job, group, and/or organization. However, compatibility or alignment is further delineated into supplementary and complementary fit. A situation in which an individual's attributes are similar to those of other individuals or to the characteristics inherent in the setting is referred to as supplementary fit. In contrast, when an individual's characteristics enhance or supply what is missing in the work environment, it is defined as complementary fit. For example, consider fit at the group level. Harmonious relations among group members might be achieved when supplementary fit on personality exists because each individual's personality is similar to the personalities of others in the work group. On the other hand, to function effectively in performing their task, a group might require a blend of people with different interpersonal strengths and personalities, for example, one

person with leadership qualities, one with skills at conflict resolution, and so forth. Thus, fit can exist when one entity provides what the other needs (i.e., complementary fit), when they share similar characteristics (i.e., supplementary fit), or both.

2. CREATION OF AND IMPORTANCE OF FIT

A number of theories explain how fit is achieved over time and why fit occurs. This section first addresses broad theories that explain the overall process that creates fit before considering current theories on why fit occurs and why it is important for both individuals and organizations.

2.1. Processes That Create Fit: The Attraction–Selection–Attrition Cycle and Socialization

Schneider's attraction–selection–attrition (ASA) cycle seeks to explain how persons and the organizational environment become naturally integrated over time. Individuals are attracted to and want to join organizations that have characteristics similar to their own, organizations restrict the range of the types of persons in them through the recruitment and selection of people who have the desired attributes and who are able to comprehend what the organization's context and goals should be, and, individuals who do not fit the organization leave, either voluntarily or involuntarily. At the same time, the kinds of persons in the organization determine the structures, processes, and general climate and culture of the organization.

The effects of the ASA cycle can be enhanced through socialization processes. Through mechanisms such as orientation programs, training, mentoring, and informal discussions, new employees learn the predominant values, attitudes, goals, social knowledge, and expected behaviors in the organization. Socializing new employees can lead to changes in their values or attitudes in the direction of the organization's values, or it can result in their leaving if the organizational characteristics continue to be incompatible with their personal attributes. Over time, through the ASA cycle and socialization, individuals within an organization tend to be homogenous and share many common personal attributes. Indeed, some research has shown that the personality characteristics of individuals within organizations are more similar to each other than those between organizations. Furthermore, individuals with longer tenure in organizations tend to have personalities that are more consistent with the climates in the organizations.

2.2. How Fit Affects Responses in Organizations

A number of different theories have been proposed to explain why fit occurs and how it affects employees' responses and reactions. One perspective is "needs–supplies." Individuals have a variety of needs, desires, and preferences (e.g., achievement, safety, nice working conditions, affiliation, pay, growth and development). When these needs are met or satisfied by the job, group, and/or organizational contexts, individuals will perceive a good fit and are likely to feel fulfilled. As a result, they will be more motivated and satisfied and will perform better. A related perspective is based on "abilities–demand," which suggests that fit occurs when an individual has the abilities required to meet what the organization demands (e.g., time, effort, commitment, skills from employees). Here, when the demands of the organization are met, employees fit the requirements of the organization, allowing for more effective functioning.

Refinements and additional explications of these broad theories have been offered to further explain why fit occurs and how fit guides behavior, reduces stress, and reduces cognitive demands. First, the similarity–attraction paradigm proposes that the perception of similarity to others in background, attitudes, and lifestyles is the most important determinant for attraction dynamics. Individuals who perceive more similarity to others are more willing to pursue further interaction, work together, share knowledge, and form productive work relationships. Second, fit purportedly helps to guide behavior. The environment of the organization is satisfying and reinforcing when it is consistent with individuals' personalities. This consistency or congruence increases the stability of behavior because employees are reinforced for their behavior, whereas incongruence is not reinforced and stimulates a change in behavior. Third, when individuals' needs are met by the organization, or when ability–demand fit is achieved, stress and anxiety are reduced, thereby allowing for greater satisfaction, attitudes, and performance. Finally, when people fit the environment, cognitive demands on employees are reduced. They can more

readily interpret the work environment correctly and discern what is valued as important. Less learning, socialization, and adjustment are required on the part of employees, allowing them to adapt to the organization more quickly and to perform more easily. Internal conflict is also likely to be reduced due to employees having more similar or better alignment of values, competencies, and attributes.

2.3. Relationships between Fit and Outcomes

Fit has been shown to play a significant role in how job applicants choose organizations and how recruiters select applicants. A higher degree of fit has also been consistently related to higher job satisfaction, increased commitment to the organization, more citizenship or helping behaviors, greater career success, reduced stress, and less likelihood of leaving the organization. Although theory suggests that positive outcomes will emerge for the group and organization as a whole, very few studies have examined the impact of fit on group or organizational effectiveness. It is important to note, however, that some P–E fit theorists argue that fit can also have a negative impact on organizations. This potential paradox of fit is addressed in a later section.

3. ASSESSING AND MEASURING FIT

A variety of different techniques, methodologies, and analytical procedures have been used in studies of fit, although debate continues with respect to which conceptualizations and procedures are superior. This section highlights some of the fundamental conceptualization and measurement issues in conducting fit research.

3.1. Commensurate Dimensions

What classes of variables constitute the P and E factors? Although this list is not exhaustive, the key variables on the P side include values, personality, needs, desires, goals, values, interests, skills, abilities, and personal resources such as time and effort. On the E side, core variables include organizational culture, values, organizational climate, values, goals, norms, skill and ability requirements, pay and rewards, and job characteristics. In studying fit, it is recommended that the P and E dimensions be commensurate or assessed in the same language and with the same classification scheme. For example, in examining P–O fit on cultural values, the P variables are often a list of 30 to 50 value statements, and individuals are asked to assess the personal importance of each value. The E variables contain the same list of 30 to 50 value statements, and assessments of how important these values are in the organization are made. Similarly, in examining the extent to which needs are met, the P factor represents a list of dimensions that individuals might desire (e.g., high pay, friendly coworkers, challenging work), whereas the E factor consists of the extent to which the work environment provides each of those same characteristics (e.g., employees receive high pay, work environment is friendly, jobs are challenging). In this way, direct comparisons of the degree of fit can be made on the same set of P and E variables or dimensions.

3.2. Direct Fit Versus Indirect Fit

Direct fit and indirect fit represent two different conceptual and operational approaches to fit. Direct fit focuses on individuals' own perceptions of the degree to which they fit into the work environment. It is derived by asking individuals the extent to which they believe they fit into the work environment, either generally (e.g., the extent to which they believe their skills fit the job) or for some particular dimension (e.g., the extent to which their values fit those of the organization). Here, for example, the focus is on how much individuals personally believe that their values fit those of the organization, but it is not known what the individuals' values or the organization's values actually are.

In contrast, indirect fit is based on separate measurements of the P and E factors, with fit being assessed statistically by jointly comparing the two P and E measures. For example, individuals' reports of their own values are compared with a separate assessment of the cultural values in the organization. In this case, the interest is in determining whether fit between individuals' values and organizational values is related to attitudes and behaviors, regardless of whether or not the individuals actually perceive this fit. Research is only beginning to address the relative importance of these different conceptualizations of fit. For example, some research has indicated that direct fit is more strongly related to individuals' behaviors and attitudes than is indirect fit.

3.3. Sources and Focus of Measurement

The source from which the P and E factors are derived can vary (e.g., individual, other employees, manager, independent assessor, objective data) depending on the purpose and research question. In many studies, both the P and E factors are measured from the same individuals. For example, individuals are asked to report on their desired characteristics of a work environment and then to report on their perceptions of their current organization's work characteristics. The degree of similarity or discrepancy between these two is taken as an index of fit. Same-source measurement is appropriate when the research question is to assess individuals' perceptions. However, the disadvantage is that relationships might be artificially inflated due to common method variance.

When attempting to address fit to the "actual" environment rather than to individuals' own perceptions of the environment, the E factors are typically derived from a separate source. For example, if the research question is to determine the extent to which an individual's values fit the cultural values of the work group, the individual might report his or her values while other members report on the values in their work group. Here, the individuals' values are compared with the aggregate of other employees' perceptions of the organization's values. Likewise, if the interest is in evaluating the effects of fit between individuals' goals and managers' goals, individuals' attributes would be compared with assessments from managers.

3.4. Analytical Issues

In testing fit, a number of different statistical indexes have been used, including difference scores (i.e., the algebraic or absolute difference between the score on the P measure and the score on the E measure), D^2 measures (i.e., the square root of the sum of squared differences between P and E across variables), and profile comparisons (i.e., the correlation between the profile of P dimensions and the profile of E dimensions). These indexes were criticized by Edwards for their conceptual and methodological shortcomings, and polynomial regression procedures have been touted as a more appropriate methodology for assessing congruence. Polynomial regression procedures allow for testing specific functional forms of congruence corresponding to a more specific conceptual model of interest. For example, one can assess whether a lack of fit results in more negative outcomes when individual values are stronger than organizational values compared with the reverse, or one can assess whether congruence when both the individual and organization values are evaluated highly is related to more satisfaction than is congruence when both the individual and organizational values are rated low. Although polynomial regression procedures have increasingly gained acceptance in fit research, this technique is still the subject of some controversy because it requires very large sample sizes and is best for examining the dimensions of fit one at a time rather than simultaneously.

4. SELECTED APPLICATIONS OF FIT

Fit has been examined across a wide range of topic areas, including selection, training, job design, motivation, leadership, goals, diversity, and teams. This section focuses on three specific applications of the fit concept in the areas of organizational culture and organizational climate, selection and socialization, and team composition and diversity.

4.1. Organizational Culture and Climate

Organizational culture and organizational climate are two key attributes that differentiate organizations from one another. Furthermore, both culture and climate focus on how employees interpret, experience, and make sense of their organizations. As a result, both culture and climate have become important elements of P–O fit.

An organization's culture is composed of its fundamental ideologies underlying assumptions and core values. It is deeply embedded in the organization and is collectively held. Climate is based on the ways in which individuals experience and describe what is happening at work. Employees' perceptions of what the organization is like in terms of its practices, policies, procedures, and routines are the foundation of psychological climate (individuals' own perceptions) and organizational climate (shared perceptions of the situation across individuals). Culture and climate are aspects of the "supplies" side of the needs–supplies theory underlying P–O fit. To the extent that these organizational supplies are congruent with and

meeting employees' individual needs, supplementary P–O fit will exist.

In terms of fit to organizational culture, much research has assessed congruence between individuals' values and the cultural values espoused in the organization. Studies have been directed at understanding the implications of both direct and indirect fit for cultural values and have demonstrated that greater value congruence is related to greater satisfaction, commitment, adjustment, performance, career success, and tenure. Furthermore, fit between employees' preferred organizational culture and their employer's actual organizational culture has been shown to predict task performance as well as to enhance organizational citizenship behaviors such as helping other employees and volunteering for tasks. Similarly, with respect to climate, some work has indicated that congruence between individuals' needs, personalities, and work environment preferences, on the one hand, and the organization's climate, on the other, is related to higher individual satisfaction at work, commitment to the organization, involvement, performance, and adjustment to work as well as to lower levels of stress, turnover intent, and absenteeism.

Unfortunately, very little empirical research has been devoted to examining the impact of P–O fit to culture and climate on group or organizational outcomes. However, some research suggests that organizational effectiveness is greater to the extent that employees' personalities overall are congruent with certain dimensions of the climate of the organization. Likewise, although theory suggests that hiring employees whose personal values are consistent with future desired organizational values can be an effective tool to promote change in the organization, research is lacking in this area.

The implication of this body of work is that individuals' well-being and quality of work life are enhanced to the extent that individuals fit, or perceive that they fit, into their organization's climate and culture. Organizations desiring to maximize these employee attitudes can foster such P–O fit through practices such as management, rewards, mentoring, hiring, and socialization.

4.2. Selection and Socialization

Organizations traditionally focused on P–J fit in their selection tactics and training and development processes because high P–J fit (i.e., abilities–demands fit) has been demonstrated to positively affect task performance, tenure, and attitudes. More recently, attention has been devoted to understanding the role of P–O fit in recruiting, selection, and socialization.

Values represent a starting point for P–O fit because an organization's cultural values affect what types of applicants are likely to seek, and find, employment there. Fit plays a significant role in how job applicants choose organizations and in how recruiters select applicants. Recruiters are more likely to make hiring recommendations to applicants who they perceive as having a stronger P–J fit, a stronger P–O fit, and values that are consistent with those of the organization. At the same time, job seekers' perceptions of value congruence are related to perceptions of their general P–O fit perceptions (i.e., how they would fit into the organization overall), which in turn positively affects attraction to the organization. Moreover, new hires whose values are more congruent with those of the organization are more likely to have higher job satisfaction and commitment to the organization and are more likely to remain in the organization 1 to 2 years later.

During socialization, new employees acquire an understanding of the attitudes, values, behavior, and knowledge required to participate as organizational members. Socialization practices, such as social activities that foster interaction, mentoring, and sets of formal and collective (group) learning experiences, can enhance alignment between newcomers' values and organizational values over time. The underlying assumption is that such tactics lead to a change in individuals' values. Although individuals' values are often thought to be relatively stable and difficult to change, there is also some evidence suggesting that employees' work-related values may be malleable if the organization's processes are strong.

4.3. Team Composition and Diversity

Fit in the group context has received the least amount of research attention to date. The distinction between supplementary fit and complementary fit appears to be particularly important when considering fit issues in teams or groups. Individuals' attitudes are likely to be more positive when the individuals fit well into the group and have attributes (e.g., skills, abilities, goals, values, personalities) that are similar to those of other group members (i.e., supplementary P–G fit). Likewise, in terms of the demographic composition of the group (e.g., age, race, sex, education), research has consistently demonstrated that individuals who are

more similar to others are more likely to have positive attitudes, better performances, and longer tenures.

However, it is also important to consider the composition of the group as a whole, that is, how the attributes of the group members complement one another to influence group functioning and group effectiveness. Effective group processes and performance may depend on supplementary fit, as in the case of a customer service team where all roles are shared by all group members and, hence, homogeneity in skills and interpersonal style may result in better team performance. Alternatively, a complement of more heterogeneous skills and styles might be needed, as in the case of retail teams dealing with diverse customers where a range of personality types may be linked to better team performance. Similarly, some research suggests that greater diversity in terms of demographic characteristics is also likely to have a positive impact on group creativity and performance. The underlying notion here is that some heterogeneity in characteristics and attributes is desirable. Complementary fit helps to ensure that all of the various skills and roles needed in the group are available.

The issue of achieving the appropriate balance between supplementary fit and complementary fit is a difficult one that has yet to be explored in detail. The understanding of the optimal profile or configuration of attributes across individuals for different types of work settings, tasks, and roles is in its infancy.

4.4. Implications

Although P–J fit can result in better task performance, research findings increasingly suggest that task performance alone is not sufficient for an organization's overall effectiveness. In particular, contextual performance or prosocial and organizational citizenship behaviors and commitment to the organization have been shown to contribute to organizational effectiveness. Furthermore, there is evidence that contextual performance and organizational commitment increase when P–O fit is high. As a result, all three types of fit should be attended to in making selection decisions and in designing practices such as training and socialization. More specifically, an organization should first consider (a) P–J fit as a match between the applicant's skills and abilities and the requirements and skill demands of the job to promote better job proficiency and performance; (b) P–O fit between the applicant's and organization's values, needs, and goals to enhance satisfaction, commitment, citizenship behaviors, and retention; and (c) interpersonal attributes and broad proficiencies of the applicant to promote appropriate compatibility and alignment among group members to enhance group performance and intragroup relations. The relative emphasis placed on each will likely differ depending on the organization's overall goals and strategies.

Some researchers suggest that P–J fit may be more important for jobs with greater technical skill requirements, whereas P–O fit may be more important when the organization has a distinctive culture. Others suggest that a multiple-hurdle approach may be the most effective method for considering fit in the selection and recruiting process. For example, one suggestion is to first focus on identifying candidates with high P–J fit and then, from those candidates, to select individuals who are higher on P–G and P–O fit. Alternative sequences might be equally plausible. For example, when the organization provides a great deal of job training to new employees, it may be more appropriate to first identify those with high P–O fit and P–G fit and then, from these candidates, to select those with stronger P–J fit.

In terms of fit to groups, the needs that are most critical to team effectiveness should be identified and prioritized based on the various roles within the group. Supplementary fit on attributes such as values, goals, and personality can help to foster cooperation among team members. Yet members of the group also need to have complementary proficiencies so that the weakness of one group member can be offset by the strengths of another so as to enhance overall group performance. Thus, teams need to identify the unique attributes (e.g., specialized skills, interpersonal proficiency, problem-solving ability) that are not prevalent in current group members and that need to be brought into the group to enhance group performance and effectiveness. Once individuals are hired, additional socialization and group relations training is likely to be important due to the diversity in skills and other characteristics among group members.

The majority of fit research has been conducted on new hires, with the underlying assumption that newcomers who do not fit their job, their group, or the organization will adjust poorly and will be more likely to leave the organization. In turn, tenured employees are expected to have greater fit. However, employees who do not fit may remain in the organization for a variety of reasons such as a lack of market opportunities and personal reasons. Furthermore, given increased competition in the marketplace, the changing nature of jobs, and increased technology, job

demands often change after an individual has joined the organization. Likewise, group cultures are likely to change as managers and employees transition in and out of the work environment. This implies that an individual's fit may decrease over time even if fit was initially high, and as such, continued attention should be devoted to fit after the hiring and socialization period. For example, training can be used to foster greater fit as job demands change, formal and informal activities (e.g., information sharing, social activities, mentoring) can aid in enhancing P–O fit, and group development interventions (e.g., continual feedback, interpersonal skills training, leader training) can further enhance fit to the group and group relations.

Finally, targeted interventions can be implemented to help employees deal with conflicts and problems that arise from diversity. Although some research shows that greater diversity on demographic variables is related to higher levels of creativity and performance for the group, greater diversity has also been related to lowered attitudes, more turnover, and poorer perceptions of group relations. To offset the potential negative impact on individuals' responses and to capitalize on the diversity of viewpoints, interventions such as training targeted at interpersonal relationships, communication, and understanding diversity may be warranted. As a result, harmonious group relations can be achieved more quickly, and people may perceive a greater degree of fit sooner, thereby enhancing both individual satisfaction and group effectiveness. Furthermore, over time and with sustained contact, the negative impact of diversity tends to dissipate, allowing positive group-level outcomes to emerge.

5. A POTENTIAL PARADOX OF FIT

The majority of research studies have consistently shown that a greater degree of fit is related to more positive outcomes for individuals (e.g., greater attraction to organizations, higher satisfaction and commitment, better adjustment, less turnover). The ASA cycle and socialization processes can create a situation whereby those who remain with an organization are a more homogenous group than are those who were initially attracted to and selected by the organization. However, several authors have argued that too much homogeneity within an organization can lead to stultification. When individuals have similar personalities, values, goals, and perspectives, new ideas and innovation may be stifled. The conformity in outlook that arises from a similarity of perspectives does not allow for alternative interpretations and makes it difficult for the organization to adapt to environmental changes. Information about the environment is incomplete, and the organization is unable to make appropriate and necessary changes to adapt to the environment.

Therefore, there is a potential paradox in fit. More positive attitudes and increased satisfaction of employees in an organization is a worthwhile goal because satisfaction is related to the quality of life for individuals as well as to greater productivity for the organization. Yet focusing on creating fit for individuals can lead to too much homogeneity in the organization, and this can stifle innovation, creativity, flexibility, and the ability to adapt, ultimately endangering the organization's competitiveness and long-term viability. How does one reconcile these conflicting findings on the relationship between fit and outcomes at different levels?

5.1. Possible Solutions

Although it is a topic of debate among fit theorists and researchers, there are several ways in which this paradox might be resolved. First, achieving a high degree of P–O fit among middle- and lower level employees can promote greater satisfaction and other outcomes for the majority of individuals within the organization. Achieving more diversity in perspectives, values, and skills among key decision makers or the top management team can foster problem analyses and consideration of alternatives so as to adapt to the environment and implement strategic initiatives as the organization moves forward.

Second, attention can be given to achieving similarity for core attributes that are critical for the organization while allowing variation for less critical attributes. Those attributes for which there should be similarity and those that may differ will vary from organization to organization depending on the organization's purpose, mission, strategy, and industry. Homogeneity in terms of the organization's key goals or most fundamental values is important so that employees perceive and understand the same goals and behave accordingly to promote the effective operating of the organization. Yet employees can vary in their attributes and perspectives around other less core goals and values. In such cases, there will be sufficient similarity to promote more positive attitudes and effective functioning but also sufficient diversity to allow new ideas and innovation to emerge.

Third, when the organization desires to change in some way (e.g., culture change, move to new markets), selection should focus on hiring people who have specific attributes that fit with the new goals. Both complementary fit (i.e., attributes that add what is needed) and supplementary fit (i.e., similarity) can be achieved simultaneously if the new attributes that are needed to effect the change are delineated and individuals with those attributes are selected while still ensuring that these individuals are similar to current employees on other attributes.

6. CONCLUSION

Insights from fit theory and research can be used to enhance organizational life at all levels: At the individual level, it can help to understand how to increase satisfaction, commitment to the organization, and task and contextual performance as well as how to raise individuals' quality of life at work by decreasing stress and anxiety. At the group level, fit can help to identify ideal team compositions based on both supplementary fit and complementary fit so as to enhance individual satisfaction as well as to promote harmonious group relations and effective functioning. Furthermore, fit can contribute to a better understanding of the processes and impact of demographic diversity, which in turn can lead to better, more effective diversity practices in organizations. Finally, at the organizational level, fit can be helpful in thinking about issues such as building a distinct culture and promoting change. It can also assist the organization in addressing issues such as turnover, absenteeism, and low productivity. In addition, P–O fit implies that organizations should shift away from the traditional focus on P–J fit in their selection processes in favor of practices geared toward hiring "whole persons."

See Also the Following Articles

Executive Development and Coaching ■ Organizational Culture and Climate ■ Personnel Selection ■ Recruitment ■ Vocational Psychology, Overview ■ Work Teams

Further Reading

Bowen, D. E., Ledford, G. E., & Nathan, B. R. (1991). Hiring for the organization, not the job. *Academy of Management Executive*, *5*, 35–51.

Chatman, J. A. (1991). Matching people and organizations: Selection and socialization in public accounting firms. *Administrative Science Quarterly*, *36*, 459–484.

Edwards, J. R. (1991). Person–job fit: A conceptual integration, literature review, and methodological critique. *International Review of Industrial and Organizational Psychology*, *6*, 283–357.

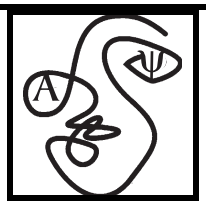
Kristof, A. L. (1996). Person–organization fit: An integrative review of its conceptualizations, measurement, and implications. *Personnel Psychology*, *49*, 1–49.

Pervin, L. A. (1989). Persons, situations, interactions: The history of a controversy and a discussion of theoretical models. *Academy of Management Review*, *14*, 350–360.

Schneider, B., Kristof-Brown, A., Goldstein, H. W., & Smith, D. B. (1997). What is this thing called fit? In N. Anderson, & P. Herriot (Eds.), *International handbook of selection and assessment* (pp. 393–412). Chichester, UK: Wiley.

Tinsley, H. E. (2000). The congruence myth: An analysis of the efficacy of the person–environment fit model. *Journal of Vocational Behavior*, *56*, 147–179.

Werbel, J. D., & Gilliland, S. W. (1999). Person–environment fit in the selection process. *Research in Personnel and Human Resources Management*, *17*, 209–243.



Personnel Psychology

David Chan

National University of Singapore, Singapore

1. Recruitment
2. Selection
3. Training
4. Performance Appraisal and Feedback
5. Compensation
6. Current Themes in Personnel Psychology
Further Reading

Personnel psychology is a subfield within industrial, work, and organizational psychology that is concerned with the application of psychological theories, methods, and research findings to the personnel functions in organizations. Primary functions include recruitment, selection, training, performance appraisal and feedback, and compensation. This article focuses on the research applications to these functions and outlines the current themes in personnel psychology.

GLOSSARY

compensation The entire process of designing and administering the package of returns that an organization provides to an employee in exchange for his or her work input.

performance appraisal The evaluation associated with a rater's evaluative judgment of the effectiveness of an employee's job performance.

performance feedback The information an employee receives concerning a rater's evaluative judgment of the employee's performance and the way in which the information is communicated.

recruitment The organizational activities and processes aimed at providing information to a pool of potential candidates to attract them to apply for job openings in the organization.

selection The process of testing and evaluating job applicants for the purpose of determining the subset of applicants to whom job offers will be made.

training The systematic and typically formal process of employee acquisition or learning of job competencies—knowledge, skills, abilities, and attitudes—that would enhance job performance.

1. RECRUITMENT

Recruitment refers to the organizational activities and processes aimed at providing information to a pool of potential candidates to attract them to apply for job openings in an organization. In the development of recruitment strategies, personnel practitioners must address important factors such as features of the relevant labor market for the job, the recruitment message and information to be communicated, and the specific recruitment methods to be used for the communication. The bulk of the recruitment research in personnel psychology is concerned with candidate perceptions of recruitment processes. These perceptions are of practical importance because candidates with negative perceptions may decide to withdraw from the application and selection process. Findings from previous studies suggest that although job attributes constitute the most important factor in affecting candidates' reactions and job choice behaviors, candidate perceptions of recruitment activities

©2004 Elsevier Inc.
All rights reserved.

and the recruiter may also affect reactions and associated behavioral intentions such as intent to apply for employment in the organization or to recommend it to others. There is some evidence that candidates tend to make inferences about the organization (e.g., leadership style) by extrapolating from perceived recruiter behaviors.

Another interest among personnel psychologists doing recruitment research is in the area of realistic job previews. The purpose of realistic job previews is to provide applicants with "true" information about the job and work environment (i.e., as the job incumbents actually experience them) rather than presenting the job and organization in "public relations" terms to maximize job and organizational attractiveness. The rationale is that the failure to present realistic job previews will lead to unrealistic, and hence subsequently unmet, expectations among newcomers (i.e., applicants who accepted the job offer) that, in turn, will have a negative impact on satisfaction, commitment, and motivation.

2. SELECTION

Personnel selection refers to the process of testing and evaluating job applicants for the purpose of determining the subset of applicants to whom job offers will be made. The majority of personnel selection research is concerned with the extent to which different selection predictors (i.e., instruments) can predict job performance and other job-relevant outcomes such as job satisfaction, organizational commitment, and intent to quit. The selection predictors that have been widely researched and used in practice include cognitive ability tests, personality measures, biodata measures, interviews, and assessment centers. Recently, there has been increasing interest in the research and use of situational judgment tests, which is one type of work simulation. Comprehensive reviews of the validity evidence of selection predictors in predicting job performance and other job-relevant outcomes were provided by Salgado and colleagues in 2001 and by Schmitt and Chan in 1998. The search for valid selection predictors needs to be driven by the nature of the criterion of interest. For example, with respect to job performance, the criterion space can be decomposed into specific areas of task requirements and specific knowledge/skill/ability requirements. In addition, the job performance criterion can be decomposed into distinct performance constructs such as typical versus maximum performance, task versus contextual performance, and routine versus adaptive performance. Each specific performance construct can, in turn, be construed in terms of multiple dimensions; for example, contextual performance can be construed in terms of interpersonal facilitation and job dedication. The validity of a selection predictor is dependent on the specific performance criterion variable in question.

In making selection decisions, it is possible to use more than one predictor, that is, to use a battery of selection tests. One major issue that needs to be addressed when using selection predictors in a test battery concerns incremental validity, which refers to how much gain in prediction power is obtained by adding the new predictor to the existing predictor(s). Issues of incremental validity have focused on effects of predictor–criterion correlations and intercorrelations among predictors. All other things equal, incremental validity is increased when the new predictor is highly correlated with the criterion but is lowly correlated with the existing predictor(s). Another major issue concerns the joint effects of the combined use of predictors on subgroup differences in overall selection score. Research has shown that adding a new predictor with a small subgroup difference (or no difference) to an existing predictor with a large subgroup difference is unlikely to lead to any substantial reduction in subgroup difference on the composite selection score.

Special topics in personnel selection include test bias (e.g., whether a test predicts similarly for applicants from different demographic groups), applicant reactions (e.g., what characteristics of the selection system would influence the degree to which applicants perceive the selection as fair), selection for staffing teams (e.g., what competencies can be tested to predict how applicants would function when working in teams), and validity changes associated with dynamic criteria (e.g., how the validity of a selection test is affected when the components of effective job performance change over time). Many of these topics were summarized by Guion in 1998 and by Schmitt and Chan in 1998.

3. TRAINING

Training refers to the systematic and typically formal process of employee acquisition or learning of job competencies—knowledge, skills, abilities, and attitudes—that would enhance job performance. The training environment may range from artificial or contrived settings (e.g., classroom lectures, role-play exercises) to highly realistic settings (e.g., on-the-job training in which the learning environment is nearly identical to

the actual work environment). Regardless of the setting, the underlying assumption is that job competencies are malleable and alterable through training. The goal of training research is to identify appropriate interventions that can positively affect employees' target attitudes and behaviors so that the employees can meet the demands of the target job. Training research typically follows the framework provided by the instructional systems design (ISD) model that emphasizes three key interrelated components of the training process: training needs assessment, training design, and training evaluation. In training needs assessment, the primary goal is to provide information on what, where, when, and who. The information gathered also serves as input for the subsequent stages of the training process, that is, training design and evaluation. The strategies to gather information may include organizational analysis to determine where training is needed; job analysis to identify the nature of the job tasks, the underlying job competencies, and the performance standards required for successful task performance; and person analysis to determine performance gaps between potential trainees' performance and expected performance levels. Training design is concerned with specifying instructional objectives, sequencing training materials, incorporating learning principles, and identifying effective training methods. There is an increasing consensus among training researchers that a training program is more effective when trainees are given the opportunity to produce the capability, when the training environment enhances trainee self-efficacy and expectations that the training will be effective in leading to acquisition of valued learning outcomes, and when performance feedback given to trainees is credible, accurate, constructive, and timely. Training researchers also appear to agree that the effectiveness of training methods is a function of the properties of the method, the nature of the training content, and the person characteristics of the trainees. For example, declarative knowledge is better acquired through classroom instruction, whereas procedural knowledge is better acquired through work simulations. However, more empirical research is needed, and advances in training research are likely to be led by rigorous studies that establish empirical interactions linking training method, training content, and trainee characteristics. Training evaluation refers to the assessment of the degree to which trainees achieved the anticipated gains in job competencies from training and effectively applied these gains to the actual job. To measure these gains and conduct an adequate training evaluation, training objectives and the intended training outcomes must be explicitly specified in ways that allow their valid measurement. An excellent overview of the typical problems in evaluating training and the alternative ways in which to assess gains was provided by Goldstein and Ford in 2002.

The ISD model, which is based on design principles derived from research findings on performance on simple tasks that require little complex cognitive processing, is somewhat deficient as a theoretical framework for training in jobs that are made up of ill-defined problems and cognitively complex tasks. Recently, several researchers have proposed more holistic approaches. These approaches involve applying cognitive constructs and principles to the training process and recognizing the multifaceted nature of the learning process by expanding the criterion space to include cognitive, skill-based, and affective learning outcomes. This cognitive-oriented and holistic context, grounded in the theory and empirical findings from cognitive psychology research, led to significant conceptual and measurement advances in training research. Many of these advances are related to the conceptualization of expertise and the measurement of metacognitive skills and knowledge structures.

4. PERFORMANCE APPRAISAL AND FEEDBACK

Performance appraisal refers to the evaluation associated with a rater's evaluative judgment of the effectiveness of an employee's job performance. Performance feedback refers to the information an employee receives concerning a rater's evaluative judgment of the employee's performance and the way in which the information is communicated.

In performance appraisal, the performance measure can be either objective or subjective. Examples of objective measures include dollars sold and number of units produced or contracts secured. Some objective measures, such as number of complaints received and days absent from work, are negatively scored. Examples of subjective measures include qualitative judgment of the creativity of work products and supervisor or peer ratings of overall job performance. The vast majority of early research on performance appraisal has focused on subjective measures and examined issues such as the various types of appraisal rating scales, rater accuracy, rater errors, and the cognitive processes underlying appraisal judgments made by raters. Despite the rater training implications of these findings, there is no clear evidence that rater

training based on these streams of research has in fact contributed to increased effectiveness of actual performance appraisal systems in organizations. One explanation for the apparent futility of early studies is that the researchers' focus on rating scales, rater accuracy and errors, and rater cognitive processes is somewhat misplaced because the researchers have neglected the central elements of the appraisal system, including the rater's motivational context and the employee's (i.e., ratee's) perspectives and reactions. This gap was filled by several subsequent streams of research focusing on motivational structures and rater–ratee relationships such as effects of rating purpose, accountability, and rater–ratee similarity on ratings and effects of organizational climate, perceived organizational support, and leader–member exchange relationships on appraisal reactions. One productive stream of research came from Greenberg's application of organizational justice research to performance appraisal in 1986. Studies by Greenberg and other researchers have identified a variety of aspects of the appraisal process and outcome (e.g., soliciting and using input prior to appraisal, ability to challenge appraisal) that affect justice perceptions about the appraisal. Most attention has been directed toward employee participation or voice in the appraisal and feedback process. In addition to identifying the antecedents of justice perceptions, researchers have examined the influence of justice perceptions on a number of attitudinal and behavioral outcomes (e.g., organizational commitment). Current research has adopted a broader perspective by locating the appraisal process in the larger context of the performance management system in an organization. This system refers to the interrelationships linking appraisal system development, appraisal processes, and feedback processes.

There has been increasing interest in "360-degree" assessment/feedback systems, where the central idea is to obtain appraisals on the same employee from multiple sources (typically supervisors, peers, and subordinates). Such systems have unique features involving multiple sources and perspectives that raise important research and practical questions regarding the impact of expanding the performance construct space on appraisal outcome and reactions. However, insofar as these systems consist of multiple rater–ratee relationships, many of the cognitive and motivational elements present in conventional appraisal are likely to remain applicable. For example, prior research findings on conventional appraisal that found that employees are more likely to react favorably to appraisal feedback when there is prior discussion on performance expectations at the beginning of the performance period, opportunity for employees to provide input to the appraisal and discuss raters' rationales for the appraisal judgments, and consideration of employees' feelings.

Performance appraisal and performance feedback have the ultimate common goal of increasing job performance effectiveness. Important conceptual and practical relationships exist between performance appraisal and performance feedback. For example, accountability effects in performance appraisal are likely to be related to the type and specificity of performance feedback information, and justice perception effects in performance feedback are likely to be related to the accuracy and transparency of the performance appraisal process. Research findings with more theoretical value and higher external validity are likely to be obtained from studies that integrate performance appraisal and performance feedback.

5. COMPENSATION

Compensation refers to the entire process pf designing and administering the package of returns that an organization provides to an employee in exchange for his or her work input. The package, which may include financial remuneration, services, and medical and other benefits, is typically divided into direct pay and indirect pay/benefits. Direct pay may consist of the fixed base pay, base pay supplements (e.g., market value adjustments, cost-of-living allowances, overtime pay), and performance-related pay (e.g., merit increments, performance bonuses, performance commissions). Indirect pay and benefits may consist of any services or benefits provided to the employee (e.g., vacations, medical benefits, discounts for goods and services).

In human resources management, most attention on compensation design and administration is focused on the strategic mix of the various components of returns to align the compensation system with the resources, constraints, priorities, and strategic goals of an organization within the labor market context of the need to attract and retain suitable employees. This typically revolves around issues of decisions concerning the value or grade level of the various jobs in the organization, the level of compensation relative to that of external referent organizations (which are typically competitors for the same employees), the design of the compensation package for individual employees, and the administration of the package. Compensation research in personnel psychology, on the other hand, is more concerned with the psychological relationships at

the individual level (although it may also involve team, organizational, or higher levels of analysis) linking compensation to employee performance and reactions. Specifically, these relationships are concerned with the motivations, perceptions, and attitudes of employees who are affected by the various types of compensation (and ways in which it is administered) and how these motivations, perceptions, and attitudes, in turn, affect employee performance and reactions.

Compared with issues of selection, training, and appraisal, compensation issues have largely been ignored by personnel psychologists, despite the fact that the idea of linking rewards to job performance has been central in empirically established theories of work motivation such as expectancy theory, goal setting theory, and equity and justice theories. Although compensation typically refers to a contractual (legal or semilegal) or formal package of returns, rewards is a more generic term used to refer to any positive reinforcement or inducement provided (explicitly or implicitly) by an organization to employees that is assumed to positively correlate with past, current, and/or future performance. During recent years, the rapidly changing nature of work has led to an increased use of new pay systems by employers. This has, in turn, activated research interest among some personnel psychologists. Not surprisingly, the majority of the recent personnel psychology research on compensation is on new forms of pay systems and has found considerable evidence that some of these new pay systems, such as variable pay and competency- and skill-based pay, are related to improved employee and organizational performance. However, it remains unclear why or how these new forms of pay lead to increased performance, although a number of conceptual models based on various psychological theories (e.g., social exchange theory, theory of planned behavior) have been proposed to account for the observed relationships. Instead of applying grand theories to commonalities across various new and effective pay systems, midrange theories that relate the specific form of pay system to the pay context, employee characteristics, and employee reactions are more likely to enhance our understanding and provide practical recommendations for specific compensation contexts.

6. CURRENT THEMES IN PERSONNEL PSYCHOLOGY

There are several discernible themes in current research in personnel psychology across the areas of recruitment, selection, training, appraisal, feedback, and compensation. One theme concerns the importance of focusing on the multidimensional and dynamic nature of job performance. In early research, job performance was construed nearly solely in terms of core technical proficiency or task performance. Current research has expanded the performance construct space to include organizational citizenship behaviors, counterproductive behaviors, and adaptive behaviors. This expansion has implications for the various areas of personnel psychology (e.g., identifying and selecting candidates with dispositions toward citizenship behaviors and against counterproductive behaviors, appraising employees on these additional performance dimensions, designing training that enhances the ability to adapt to changing situations). Current research has also focused much attention on the dynamic nature of job performance with respect to changes in performance level over time and changes in performance dimensionality over time. These changes over time raise complex but important conceptual and measurement questions that provide a springboard to productive future research. Another theme concerns person–environment fit, where environment may refer to the job, the work group, or the organization. The basic idea is that maximizing person–environment fit often (but not always) leads to positive outcomes in recruitment, selection, training, appraisal, feedback, and compensation. Research in person–environment fit brings with it several challenging methodological and data-analytic issues such as the need to distinguish between subjective fit perceptions and objective fit assessments. Finally, running across virtually all current research in personnel psychology is the central theme that the changing nature of work has major implications for the conceptualizations and measurement of the various focal constructs as well as the relationships (conceptual and empirical) between constructs.

See Also the Following Articles

Employment Interviewing ■ Industrial/Organizational Psychology, Overview ■ Occupational Psychology, Overview ■ Personal Initiative and Innovation ■ Person–Environment Fit ■ Personnel Selection ■ Recruitment

Further Reading

Chan, D. (1998). The conceptualization and analysis of change over time: An integrative approach incorporating longitudinal means and covariance structures analysis (LMACS) and multiple indicator latent growth modeling (MLGM). *Organizational Research Methods*, *1*, 421–483.

Chan, D. (2000). Understanding adaptation to changes in the work environment: Integrating individual difference and learning perspectives. *Research in Personnel and Human Resources Management*, *18*, 1–42.

Gilliland, S. W., & Chan, D. (2001). Justice in organizations: Theory, methods, and applications. In N. Anderson, D. S. Ones, H. K. Sinangil, & C. Viswesvaran (Eds.), *Handbook of industrial, work, and organizational psychology* (Vol. 2, pp. 143–165). Thousand Oaks, CA: Sage.

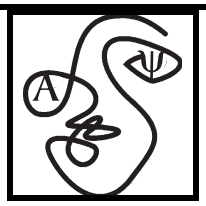
Goldstein, I. L., & Ford, J. K. (2002). *Training in organizations: Needs assessment, development, and evaluation* (4th ed.). Belmont, CA: Wadsworth.

Greenberg, J. (1986). Determinants of perceived fairness of performance evaluations. *Journal of Applied Psychology*, *71*, 340–342.

Guion, R. M. (1998). *Assessment, measurement, and prediction for personnel decisions*. Mahwah, NJ: Lawrence Erlbaum.

Salgado, J. F., Viswesvaran, C., & Ones, D. S. (2001). Predictors used for personnel selection: An overview of constructs, methods, and techniques. In N. Anderson, D. S. Ones, H. K. Sinangil, & C. Viswesvaran (Eds.), *Handbook of industrial, work, and organizational psychology* (Vol. 1, pp. 165–199). Thousand Oaks, CA: Sage.

Schmitt, N., & Chan, D. (1998). *Personnel selection: A theoretical approach*. Thousand Oaks, CA: Sage.



Personnel Selection

Deniz S. Ones
University of Minnesota, Minneapolis, Minnesota, USA

Chockalingam Viswesvaran
Florida International University, Miami, Florida, USA

1. Introduction
2. Predictors Used in Personnel Selection
3. Fairness in Selection
4. Combining Predictors
5. Emerging Issues in Selection
6. Conclusion
Further Reading

GLOSSARY

adverse impact Disproportionate hiring of different groups; in the United States and the United Kingdom, adverse impact is said to occur when the selection ratio for a legally protected group is less than four-fifths that of the majority group.

criteria Measures of how well individuals behave or perform; these are the behaviors, performances, and outcomes that the organization is interested in predicting using a selection system.

criterion-related validity The association between the predictor and criterion scores; the criterion-related validity coefficient indicates the extent to which individual differences in a particular predictor are related to individual differences in a criterion (e.g., job performance).

predictive bias The systematic over- or underprediction of a criterion.

predictor Any measured applicant trait or characteristic that is used to predict performance and other behaviors on the job; a predictor assesses applicant characteristics deemed to be job relevant using a specific measurement method (e.g., test, application blank, interview).

selection ratio Proportion of job applicants hired to those who apply for a job.

Personnel selection is concerned with identifying personal characteristics of applicants that relate to performance on jobs, measuring applicants on these characteristics, and hiring those who have the required characteristics.

1. INTRODUCTION

Personnel selection has been, and continues to be, one of the central functions of industrial, work, and organizational (IWO) psychology. Selecting the right people for the right jobs constitutes a source of competitive advantage for organizations. Individuals differ in various characteristics such as skills, abilities, personality traits, and interests. Similarly, jobs differ in what is required of the incumbents. The objective in personnel selection is to identify the personal characteristics that relate to performance on jobs, measure applicants on these characteristics, and hire those who have the required characteristics. When done properly, selection can enhance organizational productivity as well as provide individual benefits to workers themselves (e.g., increased satisfaction).

In personnel selection, one assesses the applicants on characteristics relevant to performance on the job and may want to select those who are likely to perform the best on the job. There are several characteristics on which individuals differ. Abilities, skills, personality traits, and physical abilities are some examples. These individual differences, which are relevant to

©2004 Elsevier Inc.
All rights reserved.

performance on the job under consideration, must be measured reliably and accurately before hiring decisions can be made. The method of such measurement could be an interview, a test, biographical information, or even physiological measurement.

A predictor can be any measured applicant trait or characteristic that is used to predict performance and other behaviors on the job. That is, a predictor assesses applicant characteristics deemed to be job relevant using a specific measurement method (e.g., test, application blank, interview). This article provides an overview of several commonly used and emerging predictors used for personnel selection.

Given several predictors, how should one choose the most desirable or optimal one for personnel selection in a particular situation? The answer to this question depends on the behaviors one wishes to predict. Such work behaviors can include both positive behaviors (e.g., job performance, task performance, teamwork) and negative behaviors (e.g., absenteeism, employee theft). Any work behavior that one is interested in predicting is referred to as the criterion.

Predictors assess variables on which individuals differ, and such differences are relevant only to the extent that they reflect individual differences in relevant criteria of interest (e.g., job performance, other work behaviors). Whether or not a predictor is a good prehire indicator of work behaviors that one is interested in can be determined by examining the relationship between a given predictor and work behaviors of interest. Criterion-related validity refers to the association between the predictor and criterion scores. The criterion-related validity coefficient indicates the extent to which individual differences in a particular predictor are related to individual differences in a criterion (e.g., job performance). Thus, the criterion-related validity coefficient can be used to evaluate the job relevance of a predictor for selection. Once the criterion one wishes to predict is identified, one can choose the predictor or combination of predictors with a high criterion-related validity coefficient for that criterion.

In addition to criterion-related validity, one may want to consider other factors that could affect the utility of the selection procedures. The utility of a selection system depends on the selection ratio and variability in job performance across individuals. The selection ratio is the percentage of applicants who are hired and can range from 0 to 1.0. When an organization hires all applicants (selection ratio = 1.0), the usefulness of a predictor is nonexistent because no "selection" takes place. The smaller the selection ratio, the greater the benefits obtained from personnel selection. When there is no variability in job performance across individuals, the usefulness of a selection system is also nonexistent; it does not matter who gets selected. Thus, the exact utility (i.e., value and organizational benefits) of a predictor for personnel selection depends on the context in which it is used.

Another important consideration in evaluating the various predictors used for personnel selection is applicant reactions. For example, managerial applicants might not accept that scores on a verbal ability test reflect individual differences in managerial job performance. Undoubtedly, there are also cross-national differences. For example, personnel selection in the North American context is more reflective of the organizational perspective in that research and applications have emphasized the predictive validity of selection tools. In contrast, European research not only has been more eclectic in its assessment of the predictive validity of selection tools but also has taken a multilateral approach to evaluation, including evaluations from multiple stakeholders. It is more common in the European context to view selection as a negotiation among various parties such as the organization, unions or employee associations, and governmental agencies. Perhaps this broader focus is reflected in academic discussions of social constructivism in inferring meaning of organizational and individual acts. The study of applicant reactions, although in its incipient stages, has the potential to bridge this gap between the organizational perspective and multiple stakeholder perspectives.

Applicant reactions have important consequences for an organization. If top applicants turn down job offers, the utility of the selection effort is reduced. Furthermore, dissatisfied applicants are more likely to litigate against an organizations and boycott its products. This article reviews research on the variables that affect applicant reactions and discusses different predictors used in personnel selection, both in terms of their criterion-related validity and in terms of applicant reactions.

Selection affects not only organizations and the individuals applying for jobs but also society. Society has a stake in promoting fair selection procedures, maximizing productivity, and ensuring opportunity to all individuals. The issues become complicated when there are discernible differences across demographic groups in predictor scores. Group differences result in greater proportions in hiring of the higher scoring group, and this could result in an adverse impact for the lower scoring group. Therefore, in discussing the different predictors used in personnel selection, this article also

reviews issues of fairness, bias, adverse impact, and group differences in predictor scores. Attempts made to reduce the impact of group differences (e.g., banding procedures) are also summarized. Finally, methods of using multiple measures in selection and research on emerging topics in personnel selection are presented.

2. PREDICTORS USED IN PERSONNEL SELECTION

Selection occurs after a pool of qualified applicants has been identified. Recruitment refers to the activities undertaken to attract a qualified pool of applicants, and several factors affect the effectiveness of recruitment. The content of the recruitment message, the characteristics of the recruiter, the medium used, and the like have been researched. Although this article does not elaborate on these topics, it is important to note that some prescreening (including self-selection) will have occurred in the normal flow of events in organizations prior to application of the selection methods discussed in what follows.

Several predictors are used in personnel selection. Some predictors assess whether individuals have the knowledge, skills, and abilities to perform a job task (i.e., "can do" predictors), whereas others assess whether individuals will typically perform well on the job (i.e., "will do" predictors primarily assessing motivation). In classifying predictors, it is possible to view them as determinants of declarative knowledge (i.e., knowledge of facts) needed to do the job (e.g., job knowledge tests), procedural knowledge (i.e., knowledge of how to get the job done), and motivation. A multitude of predictors have been used in personnel selection. This section provides a summary of some of the most popular predictors used in personnel selection. It organizes the predictors into three main categories: (a) predictors based on individual differences (cognitive ability, physical ability, personality, and biographical information); (b) method-based predictors that can be used to assess various knowledge, ability, and trait combinations (job simulations and work samples, assessment centers, and interviews); and (c) predictors used primarily in the prediction of undesirable work behaviors (integrity tests and drug testing).

The coverage in this article is not comprehensive given space constraints, but it is intended to familiarize the reader with some of the popular selection tools. In fact, at times, popular predictors might not be the most valid ones. Table I summarizes the use and validity of the predictors discussed in this section. The data on predictor use vary across surveys; therefore, the table indicates a general categorization of use as high, medium, and low. The validity data are culled from meta-analyses and are also summarized into broad categories (high, medium, and low). The data in the table are suggestive and not definite in any sense because averaging percentages across surveys, where the numbers of respondents differ, is not indicative of any specific trend. The table is provided only as a heuristic to enable the reader to appreciate that some organizations may be employing less valid methods and incurring large utility losses. In a way, the table suggests that society may be incurring forgone opportunity costs to the extent that popularity and validity diverge.

2.1. Cognitive Ability Tests

There are more than a dozen definitions of what constitutes cognitive ability. Common threads across most

TABLE I
Summary Status of Personnel Selection Predictors

Predictor	*Criterion-related validity*	*Extent of use*	*Applicant reactions*
Cognitive ability	High	Medium	Moderate
Physical ability	Medium	Low	Positive
Personality tests	Medium	Medium (but increasing)	Moderate
Biodata	Can be high	High	Moderate
Job simulations	High	Low	High
Assessment centers	Can be high	Low	High
Interviews	Can be high	High	Moderate
Integrity tests	High	Low (except in the United States)	Moderate
Drug testing	Medium	Low (except in the United States)	Negative

of the definitions are the capacity to process information and learn. IWO psychologists have been using cognitive ability tests for more than 80 years. Furthermore, cognitive ability is one of the most widely researched traits in personnel selection. IWO psychologists have developed several tests of cognitive ability, and meta-analytic cumulation of the vast bodies of criterion-related validity studies have shown that cognitive ability is one of the best predictors of job performance. The relationship between cognitive ability test scores and job performance averages approximately .50. This is higher than what is found for other predictors. Furthermore, the validity of cognitive ability tests has been demonstrated in nearly all situations. This predictor is perhaps the best and most applicable one in personnel selection for nearly all jobs. All jobs require some cognitive ability to learn the jobs and perform effectively. The predictive validity of cognitive ability tests tends to be higher for high-complexity jobs (.58) than for low-complexity jobs (.23). A high-complexity job (e.g., nuclear physicist) involves more processing of information than does a low-complexity job (e.g., truck driver).

In addition to tests developed to assess general cognitive ability, there are ability tests tailored for specific purposes such as assessment of accounting ability and knowledge of mathematics. Tests have also been developed to assess perceptual ability and knowledge of mechanical principles. All of these tests measure general cognitive ability (*g*) to some extent, but they also have a specific ability component. It has been hypothesized that these specific abilities could be more predictive for specific jobs. For example, it has been argued that perceptual ability may be more important in clerical tasks. However, empirical research suggests that the predictive validity of specific abilities stems from their assessment of a general information-processing ability (i.e., the general cognitive ability). If the general cognitive ability is partialled out (i.e., held constant), most of the predictive power of specific ability tests vanishes. However, specific ability tests may have more face validity and may generate positive applicant reactions in a selection context.

2.2. Physical Ability Tests

Physical ability tests have also been used in personnel selection, especially in certain jobs such as firefighting and police work. Although there is a vast empirical literature on cognitive ability, the research on physical ability in selection contexts is more limited. Researchers have identified several dimensions of physical ability. One classification conceptualizes physical ability as being composed of nine dimensions: static strength, explosive strength, dynamic strength, trunk strength, extent flexibility, dynamic flexibility, gross body coordination, gross body equilibrium, and stamina. Tests have been devised for these nine dimensions based on common physical tasks. Other researchers have categorized physical ability assessments into two major dimensions: endurance and maximum strength measures. The validity and usefulness of physical ability as a predictor of job performance has been empirically investigated. One study presented the validity of physical ability tests in personnel selection as .32 for predicting supervisory ratings.

2.3. Personality Tests

IWO psychologists have used personality variables for personnel selection since the beginning of the 20th century. Walter Dill Scott devised measures for selection of salespersons that assessed aspects of personality. The Woodworth personality datasheet was used to select or identify individuals who are likely to suffer from war neurosis. During early days, IWO psychologists hypothesized and assessed different variables that they considered to be promising. Variables such as reliability, dependability, sociability, and ego strength were assessed. Sometimes different names were given to the same trait, and at other times the same label was used to assess different traits. Given this state of affairs, the validity and usefulness of personality variables in personnel selection were questioned during the 1960s. Furthermore, the behaviorist revolution in psychology contributed to personality variables falling out of favor during this time.

Research conducted over the past decade has been more positive. Meta-analysis was used to cumulate the validity and usefulness of personality factors for personnel selection across studies. Results suggested that the dimension of conscientiousness is valid across jobs and situations, with a criterion-related validity of .23. Openness to experience was found to be correlated with the criterion of training performance, and extraversion was found to be correlated with performance in certain jobs such as sales. Emotional stability was found to be a good predictor of performance in security-related jobs (e.g., police work). Salgado summarized the validity evidence based on European samples and found that conscientiousness and emotional stability are useful predictors in Europe as well. Newly emerging research suggests that emotional stability may also be an important predictor in the United States. Although concerns have been raised that applicants could fake

their responses, the criterion-related validity is substantial for personality inventories despite the presence of any potential response distortion among job applicants.

Assessments of personality have employed structured paper-and-pencil inventories. The Hogan Personality Inventory (HPI), the California Psychological Inventory (CPI), the NEO Personality Inventory–Revised (NEO PI-R), and the Personal Characteristics Inventory (PCI) are some of the paper-and-pencil inventories. The Minnesota Multiphasic Personality Inventory (MMPI), although developed for clinical assessments of psychopathology, has also been used in personnel selection, especially in jobs where emotional stability is important for job performance (e.g., police work, firefighting).

2.4. Biographical Information (Biodata)

The use of application blanks is prevalent in personnel selection. The underlying principle in the use of application blanks is that past behavior is the best predictor of future performance. However, the question of what weights to assign to each response has been much debated. Weights reflect the relative importance assigned to each response. In general, larger weights should be assigned to responses that characterize individuals who perform better on the job. For example, if individuals with bachelor's degrees are likely to be better performers than those with high school diplomas, those applicants who check the bachelor's degree response in the application blank should receive higher scores than should those who check the high school diploma response. This is an example of how empirical scoring is done. Scoring of biodata items can be done either empirically or based on rational grounds.

In a weighted application blank (WAB), weights are assigned to several items and the scores are summed. Once a WAB has been developed, a score can be computed for each prospective applicant by weighting the response provided by that particular applicant. Sometimes a biographical information blank (BIB) could be employed. A BIB is similar to a WAB, but the former includes more experiential and personal information. An accomplishment record is a structured inventory of experiences that an individual considers to be a reflection of his or her talents.

Biodata items can be used to infer several individual differences characteristics. Decision makers can use education-related biodata items to make inferences about math and language abilities, sports-related items to make inferences about physical ability, and items on social positions held to make inferences about personality characteristics. Research indicates that biodata instruments that are transportable across organizations can be developed. That is, scores developed in one organization can be used in another organization, so it is not essential to develop biodata items specifically for each organization. Validities of .48 have been reported for such biodata instruments. In general, applicant reactions have been favorable, especially when items have focused on events under the control of individuals. Applicant reactions are more negative when invasive items that raise privacy concerns for applicants are included on biodata measures.

2.5. Job Simulations and Work Samples

One of the major criticisms of the paper-and-pencil assessment method is that it does not capture the rich, complex, and subtle ways in which a trait is manifested in the real world, especially as performance on the job. Researchers have attempted to build realism into the assessment methods by simulating as much as possible the actual conditions under which individuals have to perform their work and testing how individuals perform in those simulations. Simulations differ in the degrees of realism captured. At one extreme are simulations that have very high psychological and physical fidelity to reality. At the other extreme are assessment methods in which the individuals are presented with paper descriptions of actual situations and asked to indicate how they would respond if they encountered the situations described.

Situational judgment tests are simulations of low fidelity. As the name indicates, individuals are presented with descriptions of situations and asked to indicate their likely responses. Situational judgment tests are typically high on face validity and have been found to display high levels of predictive validity. However, it has been difficult to ascertain the traits assessed in such simulations. Work sample tests have been employed to assess both job knowledge and procedural knowledge. Work sample tests have demonstrated high criterion-related validity (averaging .54). Work sample tests are relatively costly to develop. Also, in some instances where specialized knowledge is required or where tasks pose a danger to individuals assessed or to others, it is not possible to use work sample tests in personnel selection. Applicant reactions

tend to be favorable toward work samples, probably due to high face validity.

2.6. Assessment Centers

Related to simulations, but more comprehensive in coverage of the traits assessed, are assessment centers. Assessment centers typically involve multiple assessments of applicants using various tools and methods, including tests, interviews, and several types of exercises such as leaderless group discussion, in-basket exercises, and other problem-solving approaches. Assessment centers may also include role-playing and/or oral presentations. As such, the assessment center method of job candidate evaluation can span several days. A meta-analysis of the assessment center literature has found that the average validity of assessment center studies reported in the literature is .37 for predicting overall job performance. Recent research indicates that much of this validity is derived from traits such as cognitive ability and personality characteristics assessed. The use of assessment centers has also been found to elicit positive applicant reactions due to the apparent linkages between exercises and managerial work functions. Job applicants, especially for higher complexity jobs, consider assessment centers to be more acceptable predictors of job performance.

2.7. Interviews

Interviews are by far the most widely recognized personnel selection tools used. Interviews that are structured (i.e., the same set of questions and standardized scoring system is used for all applicants) have been found to have better validity than unstructured interviews. Another issue in the research on interviews is the use of multiple interviewers or panel interviews. Researchers have explored the composition of the interview panel in terms of demographic characteristics. The demographic characteristics (e.g., race, gender) of the interviewers and interviewees, and potential interactions between the two, have also been investigated. Typically, demographic variables do not influence the interview outcomes significantly.

During recent years, IWO psychologists have introduced two new types of interviews: the situational interview and the patterned behavior description interview. The basic rationale for the two types of interviews is that the questions asked of individuals should be reflective of actual situations that people face in the real world (where one wants to predict performance). Thus, individuals are given examples of workplace situations and asked how they would respond if confronted with such situations. Validities of approximately .45 have been reported for structured interviews.

2.8. Integrity Tests

Employers have attempted to screen out applicants who are likely to engage in theft or related counterproductive behaviors. This is especially true in some industries such as retail, fast food, and banking. Physiological measures have been used to assess dependability. Since 1988, the use of polygraphs in preemployment testing has been limited to security-related jobs in the United States. Furthermore, most psychologists believe that the validity of polygraphs is doubtful. Another option is the use of paper-and-pencil tests designed to assess integrity of applicants. Some of these tests ask applicants about their thoughts about theft, their admissions of theft, and their estimations of theft prevalence rates. The rationale is that individuals who are likely to engage in theft are likely to rationalize their behaviors by reasoning that most people will engage in such behaviors. Thus, they are likely to overestimate the prevalence rates. The emphasis is on assessing applicants' attitudes toward theft. However, recent research indicates unequivocally that integrity tests assess primarily well-known personality traits of conscientiousness, agreeableness, and emotional stability. As for usefulness in prediction of work-related behaviors, meta-analytic findings suggest that integrity tests have a validity of .39 in predicting general counterproductive behaviors. The criterion-related validity for predicting actual detected theft is lower (.13), and the validity for predicting overall job performance is .41.

2.9. Drug Testing

Screening for substance abuse has increased over the years. Integrity tests can be used to screen applicants for potential theft and for counterproductive behaviors on the job, including substance abuse. However, many employers are skeptical that responses to a few items on a paper-and-pencil test will identify substance abuse. In fact, such personality assessment may be more useful for detecting proclivity for substance abuse, whereas physiological measures may be more applicable for detecting current use. Urinalysis, blood samples, and hair testing are some of the physiological measures that have been employed in testing for substance abuse. Many employers provide job offers to applicants

conditional on passing a drug-screening test. Some studies report substantial predictive validities for preemployment drug screening on a variety of job performance criteria. However, drug testing cannot distinguish between the casual or weekend user and the workplace abuser. A person who smokes one marijuana "joint" on Saturday night will have the same result as a person who is smoking marijuana continually on the job. This inability to discriminate is a major point of controversy in the drug-testing literature. Arguments have been made that personality tests, such as integrity tests, can assess the proclivity to use drugs and so can make the distinction between a casual user and a regular user.

3. FAIRNESS IN SELECTION

In discussing fairness in selection, one needs to clearly distinguish among culture-fair assessment, group differences, adverse impact, discrimination, bias, and fairness. These are related but distinct concepts. Any assessment is grounded in a cultural context to some extent. The degree to which responses are unique to a particular culture constitutes the culture-loadedness of the test. If one uses culture-loaded tests in selection, members of cultural groups for which the items were developed are likely to score better than are members of other cultural groups. However, there is a long history of test developers attempting to identify and delete such items from tests. The essential strategy is to equate individuals belonging to various cultural groups on the trait being measured and to test whether the difficulty level of items (i.e., percentage of individuals getting the items right) is significantly different across groups. If differences are found, a conclusion of differential item functioning is reached. Several statistical procedures have been developed to assess differential item functioning. However, empirical studies indicate that, at least for well-developed professional selection tests, there is no evidence of differential item functioning across racial groups.

Culture-fair assessment does not imply lack of group differences in traits. Just as there are discernible individual differences, there can be discernible group differences in traits. Groups can be defined based on race, gender, or age. When there are group differences, it is possible that there will be more members of the high-scoring group being selected. When the selection ratios (i.e., ratios of those who were selected to those who applied for a job) differ substantially across groups, adverse impact results. That is, the low-scoring group has been adversely affected by the selection procedures.

The issue of bias, in contrast to issues of fairness, is a statistical concept with a precise definition. Predictive bias is defined as the systematic over- or underprediction of a criterion. Two related concepts should be mentioned here: single-group validity and differential validity. Single-group criterion-related validity refers to the idea that the predictor is valid for one group but not for the other group. Differential validity indicates that the predictor is valid in both groups but that the magnitude of the validity differs. Finally, fairness depends on the value system of the individuals, and there are perhaps no right or wrong answers. Fairness perceptions are reflected in applicant reactions.

Applicant reactions to a personnel selection system are important for several reasons. First, applicants might not apply for jobs with an organization if they have negative reactions to the selection system implemented by that organization. Second, when applicants are offered jobs, those applicants with negative reactions might turn down the job offers. This may force the organization to offer the jobs to less qualified applicants and to increase testing and recruitment costs, all of which would lower the utility (i.e., overall value) of the selection system. Third, applicants who have negative reactions to selection procedures employed may be more likely to sue the organization for unfair discrimination. Finally, applicants who have negative experiences with an organization's selection system may boycott the organization in the marketplace and influence others to do likewise.

Of course, it is not feasible to avoid every factor that someone, somewhere, and at some time will find offensive in a selection system. Individuals who score low on a cognitive ability test are likely to blame the test. IWO psychologists have studied various factors that affect applicant reactions based on theories of organizational justice. Research has shown that applicants have positive reactions when the selection system is based on a thorough job analysis, provides multiple opportunities for applicants to demonstrate their skills, is valid, and is administered in the same manner to all applicants. Applicants care about both the outcomes (i.e., the scores obtained whether they are selected or not) and the procedures used (e.g., whether the procedures and rules are applied consistently to all job candidates). Finally, there is evidence that applicant reactions are influenced by the extent to which all applicants were treated with dignity and respect.

Empirical data have been gathered about applicant perceptions of various predictors. In general, most commercially developed predictors have been shown to be

free of bias, although group differences may exist. Organizations must consider these issues before choosing the appropriate predictors to use. When faced with group differences (even if there is no predictive bias), an organization may have to confront the potential for adverse impact for the low-scoring group when that predictor is used in personnel selection. The extent of adverse impact will depend on factors such as the magnitude of existing group differences and how selective the organization is; that is, the lower the selection ratio, the more choosy the organization is and the greater the adverse impact will be on the low-scoring group.

Organizations have tried different strategies to address this potential for adverse impact. One strategy is to specify a low score above which all applicants are considered to be equally qualified. This strategy ignores the linear relation between predictor and criterion variables for most predictors and results in substantial loss of the utility of the selection system. Another option is to search for alternate predictors where there are fewer group differences. However, alternate predictors might not capture all of the individual differences initially identified as critical for successful job performance.

Another option is to consider the use of banding. Banding refers to a set of procedures that groups some scores as not being significantly different from one another. The rationale advanced for the use of banding is the well-known fact that there is a measurement error associated with each observed score. The idea is that scores within a band all are the same; differences among them are probably due to measurement error. Bands are usually formed using the top score in a group of applicants and the standard error of measurement (SEM). Selection can be at random within the bands, or some form of preferential hiring based on other variables can be employed. However, banding is controversial in personnel selection, and several criticisms of the technique question its scientific basis. For example, the SEM is dependent on the reliability of the test. If an organization uses an unreliable test, it can generate broad bands; that is, the SEM increases with unreliable tests. Thus, an organization intent on ensuring enough selected individuals belonging to a low-scoring group can achieve its political goals by using unreliable tests.

4. COMBINING PREDICTORS

Organizations rarely rely on a single predictor for personnel selection. When individuals are assessed on several traits, how should the scores be combined? There are two options. The first is a compensatory system in which high scores on one predictor can make up (i.e., compensate) for low scores on another predictor. In a compensatory system, the scores from different predictors are combined in one of three ways. The scores may simply be added (i.e., unit weights given to predictors), different weights may be assigned to each predictor based on rational arguments, or (for optimal prediction) regression weights can be used.

The second option is a noncompensatory system in which applicants are expected to obtain a minimum score on each predictor. The minimum score is referred to as the "cut score," and extensive technical literature exists on how to set cut scores. Noncompensatory systems can be further classified as multiple-hurdle and multiple cutoff systems. In multiple-hurdle systems, applicants are assessed on the traits one trait at a time, and only those who meet the requirements on a trait are assessed on the subsequent trait. This procedure is useful when the costs of the various assessments vary. One can administer the least expensive assessments first, and the more costly assessments may be limited to only those who meet the minimum qualifications of the least costly assessments. In multiple cutoff systems, individuals are assessed on all of the traits, and those who meet the minimum cut score in all predictors are selected. Multiple cutoff systems are used when it is costly to bring in applicants for testing on each predictor.

5. EMERGING ISSUES IN SELECTION

An increased emphasis on teamwork has been witnessed during recent years. Some issues, such as whether teams should be staffed with individuals who are similar in trait levels or with individuals who complement one another in traits, need to be explored more comprehensively. The increasing globalization of business has raised the issue of expatriate selection. What variables should be considered in selecting individuals for overseas assignments? Finally, research is needed on whether organizations should select for particular jobs or for qualified applicants in general. This is important when the boundaries among jobs fade. However, such a strategy will raise a whole host of issues such as the definition of "qualified in general." The rapid advances in computer and other multimedia technology have definitely altered, if not revolutionized, how assessments are made. The role of technology is an important and evolving topic in personnel selection.

6. CONCLUSION

Personnel selection is one of the central functions of industrial–organizational psychology. This article has reviewed the various predictors used for personnel selection and how predictors are combined to form a selection system. It discussed the relationship between selection and society at large. It considered issues of fairness in selection and applicant reactions. The article underscored distinctions between compensatory and noncompensatory systems. It noted emerging issues of team selection, expatriate selection, and selection for an organization rather than for specific jobs. Personnel selection is a dynamic field in which research and practice go hand in hand in building a science around how to best select and hire individuals for jobs.

See Also the Following Articles

Achievement Tests ■ Assessment and Evaluation, Overview ■ Employment Interviewing ■ Industrial/Organizational Psychology, Overview ■ Intelligence Assessment ■ Job Analysis, Design, and Evaluation ■ Personality Assessment ■ Psychological Assessment, Standards and Guidelines for ■ Psychometric Tests ■ Psychophysiological Assessment

Further Reading

Anderson, N., & Herriot, P. (Eds.) (1997). *International handbook of selection and assessment*. New York: John Wiley.

Anderson, N., Ones, D. S., Sinangil, H. K., & Viswesvaran, C. (Eds.) (2001). *Handbook of industrial, work, and organizational psychology*. Thousand Oaks, CA: Sage.

Campbell, J. P., & Knapp, D. (Eds.) (2001). *Project A: Exploring the limits through personnel selection and classification*. Mahwah, NJ: Lawrence Erlbaum.

Cook, M. (1993). *Personnel selection and productivity* (2nd ed.). New York: John Wiley.

Eder, R. W., & Harris, M. M. (1999). *The employment interview handbook*. Thousand Oaks, CA: Sage.

Guion, R. M. (1998). *Assessment, measurement, and prediction for personnel decisions*. Mahwah, NJ: Lawrence Erlbaum.

Heneman, H. G., III, Judge, T. A., & Heneman, R. L. (2000). *Staffing organizations* (3rd ed.). Middleton, WI: Irwin McGraw–Hill.

Kehoe, J. F. (Ed.) (2000). *Managing selection in changing organizations: Human resource strategies*. San Francisco: Jossey–Bass.

Ones, D. S., Viswesvaran, C., & Schmidt, F. L. (1993). Comprehensive meta-analysis of integrity test validities: Findings and implications for personnel selection and theories of job performance. *Journal of Applied Psychology*, *78*, 679–703.

Schmidt, F. L., & Hunter, J. E. (1983). New research findings in personnel selection: Myths meet realities in the '80's. In K. Pearlman, F. L. Schmidt, & W. C. Hamner (Eds.), *Contemporary problems in personnel* (3rd ed., pp. 98–103). New York: John Wiley.

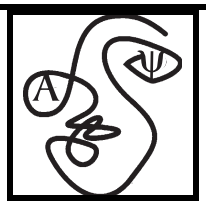
Schmidt, F. L., Ones, D. S., & Hunter, J. E. (1992). Personnel selection. *Annual Review of Psychology*, *43*, 627–670.

Schmitt, N., Borman, W. C., & Associates. (1993). *Personnel selection in organizations*. San Francisco: Jossey–Bass.

Stokes, G. S., Mumford, M. D., & Owens, W. A. (Eds.) (1994). *Biodata handbook: Theory, research, and use of biographical information in selection and performance prediction*. Palo Alto, CA: Consulting Psychologists Press.

Viswesvaran, C., & Ones, D. S. (2000). Perspectives on models of job performance. *International Journal of Selection and Assessment*, *8*(4), 216–226.

Viswesvaran, C., & Ones, D. S. (Eds.) (2002). Roles of general mental ability in industrial, work, and organizational psychology [special issue]. *Human Performance*, *15*(1/2).



Play

Cindy Dell Clark

Pennsylvania State University, University Park, Pennsylvania, USA

1. Introduction
2. Play: Developmental Processes
3. Play: Sociocultural Contexts
4. Therapeutic Play
   Further Reading

GLOSSARY

bibliotherapy Therapeutic intervention making use of books and reading.

catharsis Process by which unresolved emotional distress is relieved and resolved through expressive activity.

imaginal coping Coping through the use of imagination and play.

language play Play in which children transform and practice linguistic forms, including sounds, words, sentences, and narratives.

metacommunication Communication between players that serves to shape the play process rather than serving as the actual content of play; it serves to mutually negotiate and to maintain the play frame.

play frame Parameters for sociodramatic play, as negotiated between players; this includes which roles are to be played and which actions and objects will serve symbolic purposes in the play.

sociodramatic play Social pretend play in which children enact dramatic roles and fantasy scenarios.

zone of proximal development The gap between a child's independent ability and the ability gained when the child is helped by others who are more capable; this zone may be manifested by adult assistance in enabling a child to function on a higher level as well as through interaction with more capable peers.

Play, an activity in which children apply nonliteral flexible behavior that is motivated by pleasure rather than external goals, has been theorized to have positive developmental advantages for children's cognitive, affective, and sociocultural growth. In Euro-American society, play has been exploited for therapeutic purposes. Children make use of play for self-healing, both in their own spontaneous play and when they are guided by a professional therapist in formal play therapy. Play entails the re-representation of experience, which enables a flexible remaking of experience that has both therapeutic and normal developmental advantages.

1. INTRODUCTION

Straightforward attempts to define play have been elusive. Play, which encompasses both human and animal behavior, evades obvious definition. The difficulty traces partly to the broad and varied forms of play, ranging from rule-governed games, to improvised role-play, to rough-and-tumble mock wrestling. Play, for definitional purposes, is recognized mainly by how

©2004 Elsevier Inc.
All rights reserved.

TABLE I
Four Criteria Defining Play

Criterion	*Definition*	*Example*
Nonliterality	Takes an "as if" stance toward reality, transforming actions and meanings	The child lifts and shakes his or her head, roaring as if he or she were a lion
Intrinsic motivation	Activity is done for its own sake rather than for external motivation	The child pulls a wagon for the fun of it rather than to take things somewhere
Positive affect	Pleasure or enjoyment	The child squeals with glee when riding a swing
Process flexibility	Actor flexibly makes use of an object rather than conforming to an action dictated by the object	The child uses a large box for many varied playful uses: turning it upside down as if it were a rock to hide under, turning the opening to the side as if it were a "store," getting inside the box and rolling down a hill, and/or using it to build a pretend machine with fantasized powers

it contrasts with other dispositions or orientations. A workable approach for defining play involves assessing a behavior based on four criteria. The more the four criteria apply to the behavior, the more the behavior can be categorized as constituting play. These criteria, summarized and illustrated in Table I, are (a) nonliterality (a transformational "as if" stance such as riding on a stick as if it were a horse), (b) intrinsic motivation (an action pursued for its own sake rather than driven by external goals), (c) positive affect (pleasure), and (d) process flexibility (action taken in a free and flexible manner rather than constrained by an object). Activity that fits these criteria can be observed across the human life span.

In addition to occurring in many forms, play has ambiguous or paradoxical characteristics. For example, play is an activity set apart from ordinary reality, yet it often makes reference to ordinary social reality in its content and form. Play encompasses creative innovation by the individual, yet it concurrently is shaped by sociocultural norms. Play, an activity that is not explicitly goal directed, has also been used with therapeutic effect for problems ranging from support during hospitalization, to trauma intervention, to grief therapy among children. Despite (or perhaps in response to) the intriguing ambiguities of play, scholarly attention to play has blossomed over the transition into the 21st century.

This article discusses accumulated theories and findings about play among children, including the developmental trajectory, sociocultural context, and therapeutic application of play.

2. PLAY: DEVELOPMENTAL PROCESSES

Early in the 20th century, Groos offered the speculative theory that play served as practice for young organisms who used play to develop their survival skills. Modern-day observers of play have recognized Groos's insight as appealing. For example, when a child and a playmate pretend to go shopping for groceries, they seem to be practicing social behaviors that prepare them for actual adult roles.

Contemporary theories of play share an underlying assumption that play is beneficial and contributes positively to children's development. Influential theorists of play who wrote during the first half of the 20th century, including Freud, Piaget, and Vygotsky, all argued that play serves childhood mastery through the transformation of themes within a nonliteral domain. In other words, the representational symbolism of play helps children to master issues related to affect, cognition, language, and/or social understanding. Play is a way in which to advance children's competence and mastery and has pronounced developmental significance.

In Freud's book *Beyond the Pleasure Principle*, he argued that play is significant for children's emotional mastery. Freud asserted that play could have a cathartic effect for children, facilitating the discharge of negative feelings. Building on Freud's ideas, Erikson posited in his classic work, *Childhood and Society*, that play provides a domain in which children can find self-healing. As Erikson described, children use play to create a re-representation, that is, a model situation

flexibly manipulated to achieve emotional mastery. For example, a child who witnessed a violent crime may repeatedly reenact the scene in improvised play scenarios as if to reencounter affective issues precipitated by the violence. Play, according to Klein's *The Psychoanalysis of Children*, is a suitable medium for children to express complex ideas and feelings that they cannot yet verbalize. Therefore, for young individuals, play is a more appropriate vehicle for therapy than is conversational discourse.

Piaget focused on children's cognitive functioning in his work on play. Piaget conceived of learning as essentially a balancing of two complementary processes: assimilation (i.e., modifying reality to fit one's cognitive structures) and accommodation (i.e., changing one's cognitive structures to conform with reality). In Piaget's book *Play, Dreams, and Imitation*, he viewed play as an unbalanced state in which assimilation dominates over accommodation. Piaget believed that children achieve new learning not through play but rather through practiced and consolidated recently acquired cognitive abilities. In contrast, Vygotsky viewed play as having a more central role in cognitive development, for example, by furthering children's abstract representational thoughts through playful creation of imaginary scenarios. For example, substituting a banana for a telephone in a make-believe phone conversation would precipitate growth in children's ability to use meaning abstractly and symbolically. Vygotsky also referred to pretend play as constituting a zone of proximal development, a framed and imagined experience in which peers can perform activities together beyond a particular child's level of competence outside pretend play. Vygotsky also noted that play has affective components in that children act out desires that have not yet been fulfilled.

Much of the empirical work on children's play has focused on sociodramatic or pretend play in which children enact pretend roles and situations. Such play has been observed as early as the end of a child's 1st year of life, although true social pretend play normally emerges at around 3 years of age. In his *Play, Dreams, and Imitation,* Piaget reported an observation of his own daughter at 15 months of age where she treated a piece of cloth in the same manner as she usually treated her fringed pillow, that is, by laughing and pretending to close her eyes to signal her playful intent. Infants from varied cultures have been shown to produce pretend gestures from roughly 12 to 13 months of age. As sociodramatic play develops, children's activity gradually becomes separated from specific contexts and outcomes, contains more references to others, and becomes less reliant on realistic props or materials. For example, early pretense might entail using a cup and pretending to drink, but later pretense could involve serving a pretend beverage to a menagerie of pretend animals from an imagined cup filled with a potion having fantastic effects.

Pretend play, as studied in middle-class Euro-American children from intact families, rises in incidence from the 3rd year of life to approximately 6 or 7 years of age. Sociodramatic play declines during the late kindergarten and elementary school years. During the grade school years, another form of play, games with rules, increases.

3. PLAY: SOCIOCULTURAL CONTEXTS

Established play theorists, including Piaget and Vygotsky, made presumptions that play, its forms, and its trajectory are universal such that social pretend play might, in theory, emerge at 3 years of age in all societies. Yet children's play is cultural activity and is subject to variation within distinctive cultural milieus distinguished by particular physical, economic, and social contexts.

Play is encouraged more in some cultural settings than in others. Play covaries with economic and social factors, which may affect the opportunity to play in the face of necessary work, as well as with the overall value placed on play by a community. Goncu, who examined cultural variations in children's play in his 1999 edited book *Children's Engagement in the World,* argued that in small villages with subsistence economies, parent–child play is rare due to the pressing demands of work. Children, who are expected to contribute to the economic well-being of families in these villages, might not have time to play. In the same volume, Gaskins described Mayan children as having access to few props for pretend play. Pretend play is not assumed to be a valued activity by Mayan adults; rather, it is tolerated at best. Playful use of language, or language play, is also less frequent among Mayan children than among American children. In another study, Carlson and colleagues found that in American Mennonite society, pretend play is rarely discussed or embellished and is sometimes actively discouraged. Play themes among Old Order Mennonite children during recess adhere closely to real-life roles.

If the amount and content of pretend play is subject to cultural variation, so too are the metacommunications by which play participants interact to maintain and transform play. Martini, in an ethnographic study of peer interactions among Polynesian Marquesan 2- to 5-year-olds, reported that these children avoided lengthy one-to-one negotiations about play activities, preferring to maintain consensus by playing in ways that were familiar to all players, that is, without much change or elaboration made to play events. Marquesan children similarly avoided leadership roles by engaging in games where all participants acted together in similar ways. In contrast, American children cooperated less often so as to negotiate and individually innovate during play. In Haight's study of caregiver–child pretending, she reported on naturalistic home observations of child–mother pretend play among 1- to 4-year-olds, tracing differences in the metacommunications about play between Chinese families and U.S. families. Although both Chinese and American caregivers sanctioned pretend play and initiated pretending for purposes of socialization, their metacommunications about play diverged. Chinese caregivers initiated pretend play in a more hierarchical and didactic manner, whereas American caregivers facilitated a more mutual engagement in which children were allowed to lead the interactions. Table II offers specific examples from Haight's research.

Although developmental scholars once viewed children in social groups who engaged in relatively little pretend play as if they were disadvantaged (by the so-called "play deficit"), this view has been dismissed by culturally sensitive researchers (e.g., Schwartzman). Variations in play reflect larger differences inherent to the particular cultures of children, representing sociocultural distinctions rather than deficits. Healthy functioning in play is best understood by keeping in mind the sociocultural context of a given time and place and by assessing play within the sociocultural confines of children's own environments.

It is important to note that culture is not an all-determining factor of the form or quantity of play. Individual differences occur within a culture in children's fantasy predispositions and play styles. When society has not provided materials explicitly meant for play, children have been known to use discarded items and places (e.g., abandoned buildings, construction sites) for play. Moreover, gender-related differences have great significance, both in the form of parental discouragement of cross-gender playthings and in the form of peer interactions that reflect and shape gender differences. Thorne, in her book *Gender Play*, described how American school-age boys and girls mark boundaries on the playground (and elsewhere) between girls and boys—boundaries that are then crossed, undermined, and challenged through playful behavior. Play reflects not only individual differences and gender-related social behavior but also the distinctive cultural context.

TABLE II
Examples of Caregiver–Child Pretending: Girls and Mothers at Play

Taipei (Chinese)	*Caregiver:* Stand up, bow down. Teacher is going to deliver a lesson.
	Child: [Child smiles and moves closer to caregiver.]
	Caregiver: Teacher is coming to the classroom. What should the class monitor say?
	Child: Stand up! [Child stands and bows. Play continues as they read a story.]
Age of child: 2½ years	*Caregiver:* We have finished the story. [She claps her hands.] Before we dismiss the class, the class monitor should say 'Stand up! Bow! Sit down!' Stand up!
	Child: [Stands.]
	Caregiver: Bow!
	Child: [Bows.]
	Caregiver: Class is dismissed. Go play on the slide [She points to imagined slide.] Class is dismissed and you are happy.
Chicago (American)	*Child:* [Child gets toy, a plastic dashboard with steering wheel and seatbelt, and positions herself as if to drive.]
	Caregiver: Be sure to fasten your seat belt, right? Buckle up for safety!
	Child: [Child fastens seatbelt. Child turns to the steering wheel and pushes the horn.]
	Caregiver: Where are you going? Are you taking a trip?
Age of child: 3 years	*Child:* To Havana. [Child steers and pretends to drive.] You drive a car like this. Like this, okay? [Play continues as child explains that Havana is far and discusses with her mother how many stops will be needed en route and how old one must be to get a driver's license.]

Source. Haight, W. (1999), The pragmatics of caregiver–child pretending at home: Understanding culturally specific socialization practices. In A. Goncu (Ed.), *Children's engagement in the world: Sociocultural perspectives.* Cambridge, UK: Cambridge University Press.

4. THERAPEUTIC PLAY

Play has been regarded, both theoretically and practically, as a healing or coping activity to be engaged in when children are under stress or facing emotional challenges. A widely known application of therapeutic play takes place within the clinical modality of play therapy employed by social workers, psychologists, child life therapists, psychiatrists, and others working in counseling children. Less prominent in the scholarly literature is the fact that children also use play on their own, without explicit professional intervention, as a means of coping with adverse or stressful circumstances.

One example of child-initiated therapeutic play is the self-initiated use of stories or pretend play as a means of expression and resolution. Paley, in her well-known series of reports on young children's playful enactments, reported how children use role-playing to serve emotional ends by reworking details, for example, by eliminating the frightening big bad wolf when enacting the "Three Little Pigs," by resurrecting a dead grandfather during role-play, or by continually improvising threats to a baby (e.g., a baby-eating lion) in reflecting on the vulnerable infantile state.

Eisen, in his book *Children and Play in the Holocaust*, chronicled how children assimilated the horrors and traumas of Holocaust death camps through their own play. While waiting for deportation for internment, children pretended to comfort their dolls. Living in camps surrounded by death and cruelty, children constructed playful pastimes such as corpse tickling and pretend games of gravedigging, funerals, and gas chamber. Experiencing the most intense horrors, children found ways of playing that addressed the very traumas of their environment. Their play was a reaction to atrocity, yet it was also a buffered context in which children could creatively make meaning of a chaotic gruesome circumstance.

In Clark's book *In Sickness and in Play,* she described how American children with chronic illness used self-initiated play as a vehicle for coping with illness. Clark found that chronically ill children spontaneously, and without formal therapy, often employed play as a means of coping with the concerns and hardships of diabetes or asthma. Playfulness, such as pretending a superhero would protect the children from peril or making a game out of an onerous treatment, made space for children to examine and reframe illness-related dilemmas.

In short, play need not be guided by an adult (therapist, teacher, or child life specialist) to provide meaningful transformative value to children confronting everyday stresses. Even without clinical intervention, play seems to be a child-manageable domain for raising and addressing emotional troubles. Play allows children to approach and express feelings that, if not masked in the "disguise" of play, might be inexpressible, whether through independent play or in therapy.

The development of formalized play therapy began during the 1920s, when Anna Freud advocated the use of toys and games as a way in which to build relationships with child patients. She claimed that this method was superior to the verbal adult method of free association. Her influential idea was that whereas adults "talk out" conflicted emotions, children are better suited to "play them out." Her techniques were further developed by Klein, Axline, Erikson, and others. Klein proposed that play was children's natural means of expression and could serve as a direct substitute for verbalizations, providing sustained symbolic expression to be interpreted by analysts. Axline, inspired by the client-centered therapy movement, proffered that in therapeutic play children should be given free rein to express feelings without judgment or direction by therapists. Erikson regarded play as a natural self-healing measure, that is, a given ability of childhood by which children could be helped to help themselves. Play therapy has developed in multiple directions during recent years, serving as an adjunct to other modes of treatment such as behavior modification and proliferating in forms such as art therapy, sand tray therapy, storytelling, music and dance therapy, puppet therapy, costume therapy, bibliotherapy, and numerous forms of symbolic and game play.

Play therapy has been used in individual therapy one child at a time, but it has also found wide application with groups of children who are facing or have experienced a shared trauma such as war. As an example of the latter approach, Simo-Algado and colleagues used therapeutic play to strengthen personal resources among children in Kosovo who were exposed to trauma during Serbian occupation and the ensuing 1999 NATO military campaign. The goal of the intervention was to preventively buffer against further effects of trauma among a child population that was impoverished, subject to injury from land mines, and dealing with difficult issues such as missing persons, the discovery of mass graves, and the memories of massacres, rapes, burnings, and shootings. Under circumstances of war, the play of Kosovo children had been narrowed to the play theme of reenacted war. Occupational therapists and aids worked with the children, engaging them in other activities that reflected their shared

cultural identity, for example, songs and poems in Albanian and games incorporating native cultural themes. Drawing, clay, narratives, and puppets were introduced as media through which children could vent, process, and discuss their feelings, including fear, anxiety, hatred, and despair. Play, which indigenous teachers were trained to facilitate, provided support, insight, external expression, and a sense of mastery. Following the therapeutic sessions, children took part in sports such as soccer, providing an opportunity for adults to observe the children at play (as well as those children who seem inhibited in free play) and to identify particular children who might benefit from additional individual attention. Through the play-based interventions, children gave expression to their war experiences and then gradually reflected new meanings about past trauma, in the words of Simo-Algado, "moving from victims to survivors." The United Nations Children's Fund (UNICEF) also used art and play in its psychosocial assistance program for children in the former Yugoslavia, through which children were encouraged to recall events from the past and express hope for the future. In crisis circumstances of mass trauma, play interventions can be taught to local professionals and volunteers, who in turn can help many children in need of healing.

Play therapy has also become a mainstream means of clinical intervention with children in individual therapy. For example, play is used as a means of helping pediatric patients by child life specialists who use play to provide child-relevant support in hospitals. Play is implemented as a means of instructing and preparing children for medical treatment procedures as well as a means of expressing or relieving emotional distress among ill youngsters.

Play serves multiple therapeutic purposes. First, it is an appropriate vehicle of communication between a professional and a child. Second, play can serve to relieve emotional tension by providing a means of expressing and reworking issues. Finally, play therapy can introduce children to new forms of beneficial play behavior, potentially to be extended into children's lives outside of therapy.

Although play need not be guided by an adult (therapist, teacher, or child life specialist) to provide meaningful transformative value to children confronting everyday stresses, adult support of therapeutic play is helpful when children face extraordinary circumstances. With or without clinical intervention, however, play seems to be a child-manageable domain for raising and addressing emotional troubles.

Nevertheless, the underlying process by which emotional resolution takes place in play is not completely understood. The therapeutic utility of play and the role of sociocultural contexts in shaping the trajectories of play are areas in need of additional investigation. Although play has come to be appreciated as a powerful human activity, its ambiguity and transformative dynamism continue to pose challenges for further scholarly understanding.

See Also the Following Articles

Child Development and Culture ■ Coping

Further Reading

Bergen, D. (Ed.) (1987). *Play as a medium for learning and development*. Portsmouth, NH: Heinemann.

Bruner, J., Jolly, A., & Sylva, K. (Eds.) (1976). *Play: Its role in development and evolution*. New York: Basic Books.

Clark, C. D. (2003). *In sickness and in play: Children coping with chronic illness*. New Brunswick, NJ: Rutgers University Press.

Goncu, A. (Ed.) (1999). *Children's engagement in the world: Sociocultural perspectives*. Cambridge, UK: Cambridge University Press.

Oremland, E. (2000). *Protecting the emotional development of the ill child*. Madison, CT: Psychosocial Press.

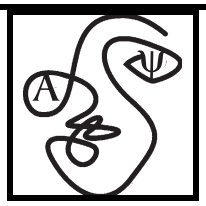
Schwartzman, H. (1978). *Transformations: The anthropology of children's play*. New York: Plenum.

Simo-Algado, S. (2002). Occupational therapy intervention with children survivors of war. *Canadian Journal of Occupational Therapy*, *69*, 205–217.

Sutton-Smith, B. (1997). *The ambiguity of play*. Cambridge, MA: Harvard University Press.

Thorne, B. (1994). *Gender play: Girls and boys in school*. New Brunswick, NJ: Rutgers University Press.

Webb, N. B. (1999). *Play therapy with children in crisis*, 2nd ed.: *Individual, group, and family treatment*. New York: Guilford.



Political Psychology, Overview

Patrizia Catellani
Catholic University of Milan, Milan, Italy

1. Scope of Political Psychology
2. Political Knowledge and Political Attitudes
3. Political Identity and Conflict
4. Political Communication
5. Political Participation
Further Reading

GLOSSARY

accessibility The ease with which information in long-term memory can be retrieved.
bias A systematic distortion in the processing of information.
entitativity The extent to which a social category is perceived as having real existence.
heuristic A simplified procedure or rule of thumb for making a decision or solving a problem without systematic processing of information.
ideology-based voting Voting choice based on ideology or values.
issue-based voting Voting choice based on parties' position on given political issues.
matching effect The different impact of persuasive messages as a function of authoritarianism and other characteristics of the target.
media agenda setting Decision process regarding the relative space and time to be devoted to different news by the media.
performance-based voting Voting choice based on evaluation of past performance of a party or a political leader.
political identity The part of the self-concept that derives from belonging to a political group.
salience Prominence, distinctiveness; any aspect of a stimulus that causes it to stand out and attract attention.
third person effect A tendency to overestimate media influence on others and underestimate media influence on oneself.
trait-based voting Voting choice based on personal characteristics of a political candidate that may have either a positive or a negative connotation.
voter's paradox A definition of voting as an apparently contradictory action, since costs implied by the act of voting are higher than the benefits the single citizen may expect from that act.

Political psychology investigates representations and actions of political actors, be they citizens, politicians, or members of groups characterized by collective and public goals. This overview begins with a definition of the area of interest of political psychology and of the main stages that have characterized the development of the discipline. We will then examine how citizens know and evaluate political reality, with a special focus on political candidates. Attention will also be devoted to the differences between experts and novices in the political domain, and superordinate principles (such as values and ideologies) around which political attitudes may be organized. In the third section of the article, political identity will be dealt with. Different groups one may identify with in the political context will be highlighted, with a particular stress on regional, national, and supranational identities, and on the roles they play in international conflicts. Media and various

©2004 Elsevier Inc.
All rights reserved.

forms of political communication, such as political interviews or discourses, influence most of our interpretations of political reality. These issues will be analyzed in the fourth section, with special emphasis on motivations and goals that underlie political messages, as well as their persuasive effects. Finally, we will consider psychological factors involved in voting behavior and in other forms of political participation.

1. SCOPE OF POLITICAL PSYCHOLOGY

1.1. Area of Interest

Let us consider millions of people in different countries all over the world gathering on the same day in main squares and streets of their towns and demonstrating in favor of peace, shouting slogans against terrorism and war. Let us consider other millions of people watching the event on television. If we zoom in on any of the demonstrators, politicians, bystanders, or the people simply watching this political event, we may investigate how they represent people struggling or not struggling for peace, what expectancies they have regarding possible success of peace movements, what plans they have regarding their future engagement in peace movements, and so on. We may also move beyond the specific event and investigate their political views, their political identities, or their voting intentions. All these issues fall within the area of interest of political psychology, which studies the representations and actions of all the people involved in political reality, from simple citizens to political leaders.

Most research in political psychology originates from the extension of "basic" psychological research to the political context, an extension that implies consideration of two main categories of context-specific factors: (1) rules and constraints that are present in the political context, and (2) motivations and goals of individuals and groups who act in the political context. For example, research on how citizens form an impression of political candidates takes into account the fact that in the political context, unlike other contexts, knowledge of a target person is often not direct, but is filtered by media. It also considers that motivations underlying the citizen's perception of a political candidate may be different from those underlying perception of a target person in other contexts, and these differences may have an influence on perception accuracy or on traits perceived as most relevant. The influence of the citizen's ideological orientation on candidate perception is also taken into account—a variable that may not be as relevant in other contexts (see Section 2). These are examples of how research in political psychology may identify a series of factors that moderate or mediate effects already observed by basic psychological research, in some cases also highlighting previously neglected effects.

1.2. Historical Development

A short reference to the historical development of political psychology may offer an idea of the main issues dealt with by the discipline so far, as well as an idea of the psychological theories that have been most frequently assumed as frames of reference. Four main stages may be identified.

1. *1940s to 1950s: Politicians' personality*. Studies of political psychology initially focused on the personality of political leaders, assuming psychoanalysis as their main frame of reference. Biographies of famous leaders (e.g., Alexander the Great, Hitler) were carefully reconstructed through historical and archival data, with a special stress on the educational environment that characterized the leaders' development. Features of the adult personality, for example, a strong tendency to dominate over others or to be aggressive toward external enemies, were traced back to a series of psychological processes originating in childhood. Other studies, not focusing on political elites but on the wider category of political activists, adopted Maslow's motivational model as a frame of reference. According to this model, people would perceive the need for engaging in political activity only when other, more basic needs have already been satisfied (e.g., health or social needs). As a consequence, politically active people would be characterized by high self-esteem and strong personal efficacy, that is, they would be deeply convinced of being able to engage successfully in political action.

2. *1960s to 1970s: Public opinion and voting behavior*. In this period, research focus changed dramatically from political elites to normal citizens and from personality to behavior. Surveys aimed at tapping political attitudes and voting intentions were being carried out on a regular basis, before or after national political elections. In 1960, Campbell et al. published a volume, *The American Voter*, in which they suggested that political attitudes would develop quite early, influenced by the individual's family environment and, by adult age,

would become rather stable and resistant to change. Party choice would therefore be the consequence of a largely irrational process, defined as party identification, instead of being based on a thorough consideration of information regarding candidates, parties, and political issues. Although this approach evokes the image of an irrational voter, a very different, substantially opposite approach to voting behavior also developed in the same years, this one supporting the image of a rational voter. In *The Economic Theory of Democracy* (1957), Downs made reference to rational choice theory (RCT), according to which individuals, in this case voters, may effectively use all the information they have at their disposal so as to choose, from a series of possible alternatives, the one characterized by the maximum subjective expected utility. Decision-making would consist in an accurate evaluation of the pros and cons of each alternative in order to choose the one that is perceived as the potentially most useful for the individual, where utility is intended as functional to the personal, mainly economic, interest of the individual.

3. *1980s: Political cognition*. As compared to the previous stage of research in political psychology, in this stage attention shifted from political attitudes and behaviors per se to the cognitive processes that underlie and influence attitudes and behaviors. The 1980s were characterized by the extension of the "social cognition" approach to the political domain. Political psychology borrowed from social cognition a well-articulated corpus of theory and data regarding what happens in the different stages of human information processing, from initial coding of new information, to its organization in the individual's mind, to its retrieval from memory. The citizen was seen as actively selecting and processing political information, attributing meaning to it through comparison with information already stored in memory, thus arriving at a personal interpretation of political reality. Information processing is constrained by the limits of human memory, mainly the capacity for dealing with only a limited amount of information at the same time. Because of these limits, citizens often do not thoroughly examine all the information they are exposed to. To save mental energy, they recur to simplified and rather superficial reasoning strategies, so-called heuristics. These strategies are functional to the everyday goal of making decisions and planning behaviors in a reasonably short time, but, being based on partial and approximate processing of available information, they may also lead to systematically distorted evaluations (biases). Thus, the political cognition perspective offers an alternative to the above-mentioned opposition between the irrational and rational voter by suggesting that voters are characterized by a limited rationality.

4. *1990s to today: Social psychology of politics*. Although the synthesis of research developments that are still unfolding should be made with some caution, two main tendencies in contemporary political psychology may be singled out. The first one involves an enhanced attention to context-related factors (citizens' motivations and goals, power relationships, normative and political constraints, etc.) and how these factors may influence both quality and quantity of political information processing. For example, cognitive theories of decision-making may be applied to voting decisions, taking into account the fact that context-related factors, such as the voting system of a given country or the dimensions and numbers of the opposing parties, may condition the way in which citizens make their decisions. The second tendency identifiable in contemporary political psychology is a gradual abandoning of a substantially individualistic perspective and the assumption of a more marked social perspective. Considering that one of the functions of politics is to regulate the relationships between individuals and groups, especially resources and power distribution, it is difficult to conceive of a psychology of politics that is completely detached from the social dimension. Social identity, social influence, intergroup relations: these are some of the factors that have been widely investigated by social psychology and are now also being taken into account by political psychology.

2. POLITICAL KNOWLEDGE AND POLITICAL ATTITUDES

In this section, we consider how political reality is perceived, felt, and reflected upon by the individual. As we have seen, studies of political psychology in the 1960s focused on public opinion: the substance of people's political attitudes and how their attitudes on various political questions are interrelated. While these questions remain relevant, and indeed central to political psychology, in the following years there was greater recognition of the interdependence of attitudes and cognition, namely, how being in favor of or against a political candidate or issue is related to what we know (or believe we know) about them.

As compared to the past, contemporary society has developed an amazing number of ways through which information, including political information, can be transmitted and disseminated (TV, radio, Internet, mobile phones, etc.), and the potentiality of these media is still far from having been entirely explored. This means that, in principle, everybody has access to an incredibly large amount of information. However, such an access is filtered and regulated by a number of selection processes. Whereas some of these processes are independent from the individual, others depend (at least partly) on them.

1. *Media agenda setting*. This information selection process is operated by the media, that is, by those who decide the relative space and time to be devoted to each piece of information, as well as the language and images that convey information. These choices inevitably reflect a given way of reconstructing and interpreting reality, and are not under the control of the individual who receives information via the media.

2. *Non-voluntary media exposure*. This second source of selection is linked to the individual's life but at the same time is largely not dependent on the individual. Being born in a given place, living in a given family environment, attending a given school, these are all factors that inevitably influence the kind of media, and therefore the political information, to which the individual is exposed.

3. *Voluntary media exposure*. Although influenced by the individual's social and family sphere, media exposure is also voluntary, because individuals, once they have reached a certain developmental stage, start to choose the kind of media to which they want to be exposed.

4. *Selective attention*. Being exposed to a given piece of information does not necessarily mean paying attention to it. Social cognition studies have clearly shown that people engage in a process of selective attention, more or less consciously, that leads them to process only the information that, for various reasons, they perceive as salient or relevant in the given context.

In this section, we shall examine the latter type of selection process. We shall see what information is more likely to catch the individual's attention, with a special focus on information regarding political candidates. We will then consider how a well-established attitude regarding a political candidate or issue may influence or bias the processing of new information. Differences between experts and novices in political information processing will also be highlighted. Finally, we shall consider various theoretical proposals regarding organizing principles (ideologies, values, etc.) of political knowledge and attitudes.

2.1. Candidate Perception and Evaluation

Citizens process information regarding political events, issues, programs, parties, candidates, or leaders. These various categories of information are obviously related to each other. However, the relative salience of each category may vary according to the goal pursued in processing that information and other context-related factors. In particular, when elections are upcoming, information regarding political candidates is especially likely to become salient. Actually, perception and evaluation of political candidates has been the most investigated area of political information processing so far, borrowing from the consistent body of results gathered by social cognition research regarding how people form descriptive and evaluative judgments about target persons.

2.1.1. Traits

When we form a judgment about a political candidate, we are likely to devote our attention to three main information categories: (1) the candidate's party, (2) the candidate's position on political issues, and (3) the candidate's traits, that is, personal characteristics that may have either a positive or a negative connotation. An analysis of the answers to an open question in a pre-electoral survey carried out on a regular basis in the United States (National Election Study) has shown that when people are asked to speak about candidates, they focus on their personal characteristics more than on their party or their position on political issues. This tendency is not significantly influenced by the individual's education level, suggesting that the stress on personal traits is not necessarily due to a lack of political competence, but may instead be due to the prevailing need for the individual to choose the right person to represent his or her own political needs and positions. Although some variation in the relative stress on personal traits has been found according to the political system and context in which individuals find themselves, recent years have witnessed an increasing personalization of politics, which is likely to make perception of the candidates' traits more and more important in citizens' political choices.

There is substantial consensus on what traits an individual focuses on when evaluating a political

candidate. These traits may be divided into five main categories: (1) competence (intelligence, ability in dealing with political issues); (2) integrity (being honest or, on the contrary, being prone to corruption); (3) reliability (being consistent, being able to keep one's word); (4) charisma (having capacity for leadership); and (5) image (being good-looking, attractive, photogenic). While the importance of the first two categories is generally more heavily stressed than the others, the relative weight attributed to each category may vary according to the context in which candidate evaluation takes place. For example, the salience of traits related to the candidate's integrity may be enhanced in a context characterized by a series of scandals involving corrupt politicians. These events usually get wide media coverage and may induce citizens to construct their evaluation of candidates around the dimension of integrity versus corruption. In these cases, a "figure-ground phenomenon" is likely to happen, in that the honesty of a given politician becomes especially salient because it is perceived as peculiar in a context of diffused corruption.

2.1.2. Emotions

The evaluation of candidates is influenced not only by their traits but also by the emotions (happiness, anger, etc.) they trigger. We may in fact wonder whether our judgment of a given candidate is based more on the traits we attribute to him or her (cognitive component of the attitude) or on the emotions we associate with him or her (affective component of the attitude). Research suggests that emotions, especially positive ones, are very good predictors of candidates' evaluation and voting intention. However, the relative weight of traits and emotions can vary according to the context. It has been suggested that emotions might become especially relevant in a changing context, when new candidates or new parties emerge in the political scene. The relative weight of positive versus negative emotions is also likely to depend on the context. Consistent with the figure-ground phenomenon described previously, higher relevance of positive emotions might be due to the fact that citizens are especially inclined to notice something positive when the context is mainly negative.

2.1.3. Influence of Ideological Orientation

So far, we have focused on traits and emotions that are generally perceived as relevant when evaluating a political candidate. However, some differences in perceived relevance of traits and emotions have also been highlighted. In particular, citizens with opposite ideological orientations have been shown to differ in the positive traits they perceive as most relevant in political candidates. In a study by Caprara *et al.*, left-wing citizens asked to rate personality traits of candidates of different parties stressed the dimension of sincerity/reliability (sincere, loyal, reliable, accountable, authentic, etc.) in candidates of their own political side and the absence of the same traits in candidates of the opposite side. Right-wing citizens performing the same task have instead focused on the dimension of energy/innovativeness (dynamic, resourceful, resolute, energetic, creative, etc.), rating these traits as high in candidates of their own side and as low in candidates of the opposite side.

Interestingly, when evaluating their own personality, citizens tend to rate themselves as high in the same traits they perceive as typical of their favorite candidates. Can we therefore infer that citizens tend to choose candidates that are similar to them? Not necessarily. It may well be that an assimilation effect is taking place here, according to which we are inclined to perceive people we like as more similar to ourselves than they actually are.

2.2. Assimilation and Contrast Effects

As social cognition research has widely demonstrated, once a positive (or negative) attitude toward a given target has developed, processing of new information regarding that target may be biased, in the sense that we look for a confirmation of our own attitudes. This may happen in various ways, through discarding inconsistent information, devoting selective attention to consistent information, or reinterpreting neutral information in a way that is consistent with our attitudes. Attitude-consistent processing of new information is especially likely when the relevant attitude is strong and when information presents itself as somehow contradictory or ambiguous. Both conditions may be often found in the political domain. People typically develop strong attitudes in favor of or against political candidates, parties, or issues. At the same time, these attitude objects are not always stable over time, but may instead change, sometimes rather substantially. Let us consider a politician who argues about the opportunity for a country to restrict or enlarge its immigration policy. The politician's position on such an issue is not necessarily the same in different places and at

different times. With regard to place, the politician may present his or her stand on the issue with different tones and nuances according to the audience. Politicians know very well how to tailor their message, using all the possibilities offered by the richness of language and rhetorical strategies (see Section 4). Regarding time, politicians may change their position on a given issue due to several reasons, such as shifts in political alliances or changes in the national or international political situation. The more mutable and ambiguous a political position on an issue, the more strongly the citizen may be inclined to select, interpret, or even distort information in a way that is consistent with his or her consolidated attitude.

If we have a positive attitude toward a candidate, we are especially inclined to perceive the candidate's position on relevant political issues as being very similar to our own (assimilation effect). For example, we may ask people to rate their position on a restricted immigration policy on a 7-point scale, ranging from 1 (fully against) to 7 (fully in favor). Then we may ask them to use the same scale to rate the supposed position of political candidates on the same issue. Research results show the presence of a high and positive correlation between one's position and the one attributed to the preferred candidate: if one's position is rated with a 5, the candidate's position will also be rated with a 5, whereas if one's position is more extreme (i.e., rating 7), the same extreme rating will be attributed to the candidate. An opposite tendency to differentiate our position on a given issue from that of a candidate we do not like has also been observed (contrast effect). The contrast effect is weaker, however, than the corresponding assimilation effect. This asymmetry has been attributed to a positivity effect, according to which people prefer to focus their attention more on what they like than on what they dislike.

2.3. Expertise

So far, we have focused on how people know and evaluate political candidates and issues irrespective of the degree of competence these people have in the political domain. However, political information processing appears to vary according to people's level of expertise. Political expertise has been defined as the synthesis of different dimensions, such as political interest, political knowledge, and media use. Experts and novices in the political domain seem to process information very differently, and these differences affect all stages of information processing.

1. *Coding*. Experts understand new political information more quickly than novices do, and are quicker in expressing evaluations on political issues and candidates. Such evaluations are also more stable over time. These differences are due to the fact that frequent reasoning about politics makes a number of political concepts immediately accessible to the experts' minds and therefore easy to employ in the interpretation and evaluation of new information.

2. *Organization*. Experts' political attitudes are more consistent than those of novices, and experts experience a larger number of thoughts related to a political candidate or issue. This difference has been attributed to how concepts and their positive or negative connotations are organized in people's minds. Experts' conceptual networks are not only characterized by a larger number of concepts, but also by a larger number of links between concepts. Moreover, a hierarchical structure is often well developed in experts' mind, so that a series of concepts is organized around a limited number of more abstract concepts. This means that the activation of one concept easily leads to the activation of other related concepts at the same or a higher level of abstraction. Such an interrelated network increases the experts' consistency. Different political issues, such as unemployment subsidies, norms toward immigration, and the welfare state, may be linked to each other in the experts' mental network, and all of them may in turn be linked to a more abstract concept such as solidarity. If the experts' attitude toward applying the solidarity principle in politics is positive, related attitudes will also be positive. This may not be the case for novices, who may be more prone to express inconsistent attitudes. At the same time, novices are likely to be less sensitive to inconsistency in politicians' attitude.

3. *Retrieval*. Better organization of political information in experts' minds facilitates the process of retrieving information; this explains why the experts' recall of political information is better than that of novices.

2.4. Ideological Orientation and Values

As we have seen, specific political concepts and attitudes may be cogently linked to more abstract, unifying principles. But what are these principles? In his article in this encyclopedia entitled "Ideological Orientation and Values," Hans De Witte identifies them with values and ideologies, showing how research on this issue has developed over time. Research on values has focused

on single relevant value dimensions, such as conformity versus self-direction or materialism versus post-materialism, but it has also developed more encompassing models, including several value types (e.g., power, security, conformity, or universalism) and their compatibility relations. Prominence attributed by individuals to one or the other of these values has been shown to be a good predictor of political attitudes and voting. Research on ideology has especially focused on the notion of conservatism (as opposed to progressiveness) and has highlighted two different dimensions of it, economic conservatism and sociocultural conservatism. Like values, ideology dimensions are predictive of political attitudes and voting.

In his article "Human Rights," Willem Doise also makes wide reference to values and ideology. The general relevance people attribute to human rights is related to values, and this reaches its highest level in people for whom the values of universalism and social harmony are very important. People may also differ in their representation of human rights according to their values and ideology. For example, for some people, the right to equality means that everybody has to be treated in the same way by the law or by public institutions, whereas for others, equality also implies that all should have equal access to a decent standard of living, education, and health care. Finally, people with different values and ideologies seem to differ in their opinions of the relative role individuals and institutions (e.g., governments) should play in safeguarding human rights.

3. POLITICAL IDENTITY AND CONFLICT

Campbell's early reference to the notion of party identification as a way of explaining people's voting behavior (see Section 1) undoubtedly had the merit of highlighting how through their vote people may express a feeling of belonging to a given party—a feeling made up of a mixture of cognitive and emotional components. However, the notion of identification as developed by Campbell had a limited heuristic validity. He saw strong party identification as a cause of a stable vote for a party, but envisaged no specific measure of identification. On the contrary, it was precisely the observed stability of voting that induced Campbell to infer the presence of identification. As such, his explanation of stable voting through identification amounts to little less than a tautology. In addition, Campbell's view of identification as being deeply rooted in developmental age failed to explain possible shifts in voting preferences.

Research on identification processes in the political domain has gathered new strength in recent years. Thanks to the substantial progress made by psychosocial research on identity, we know a lot about how people define themselves as individuals or as members of meaningful groups, how they perceive members of other groups, and what processes characterize intergroup relations, leading to overt competition and conflict or, on the contrary, to peaceful coexistence and cooperation. While extension of this approach from the social to the political domain has proved to be fruitful, specific research in this field has also started to highlight some peculiarities of identification processes in the political domain.

3.1. Political Identity

In some circumstances of our lives, we tend to define ourselves as unique and different from anyone else (personal identity). In other circumstances, we define ourselves as members of a group: we perceive ourselves as similar to the other members of the group, sharing the same values and goals (social identity). One may identify with several groups (multiple identities); these may include groups that one cannot choose to belong to (e.g., gender or race) or groups chosen by the individual (e.g., a sports team, an association, a circle of friends). Moreover, some of these groups may be nested, with a more inclusive or superordinate group (e.g., the company one works for) including other less inclusive or subordinate groups (e.g., a specific sector of the company or a specific workgroup). According to the context, a given identity may be perceived as more salient than others. For example, party identity may become especially salient during the electoral period, when people discuss a politically sensitive issue, or when we confront ourselves with people who belong to different parties. Apart from the context, a given identity may also be more "chronically" salient to the individual; this means that a given identity may be especially relevant to the individual's self-definition. As compared to other kinds of identity, political identity seems to be especially characterized by the need for expressing the group's values in public and, more generally, for acting together in order to gain further consensus around these values. As Bert Klandermans highlights in his article "Collective Action," a shared identity is one of the main antecedents of people's

involvement in collective action. At the same time, being involved in collective action leads to enhanced political identity.

An analysis of autobiographical accounts of political activists carried out by Catellani *et al.* has shown the presence of three main levels of political identity, characterized by a different degree of abstractedness and embedded within each other: ideological identity, party identity, and sub-party identity. These identity levels are likely to satisfy different needs, and, accordingly, the relative salience each of them assumes for the individual is likely to vary depending on circumstances. Ideological identity is characterized by the highest degree of abstractedness. When people identify themselves with left-wing or right-wing people, they state their inclusion within a social category that is difficult to visualize, with wide and ill-defined boundaries, including a large number of heterogeneous individuals. At the same time, such an abstract category of identification is likely to be perceived as more stable over time and less subject to transformations as compared to a more concrete category, such as a political party. In fact, ideological identity has been shown to be especially salient after a political schism, when a previous existing party splits into two or more new parties. In these changed circumstances, stressing one's ideological identity has the function of safeguarding a certain degree of continuity with the past, to keep a balance in one's own identity.

3.2. Regional, National, and Supranational Identity

Investigation of how multiple or nested identities may exist and coexist in the political domain may be usefully extended to the analysis of regional, national, and supranational identities. The issue has become of increasing relevance in a global world that is witnessing an ever-increasing close cohabitation of citizens of different nationalities as well as attempts to create solid supranational realities such as the European Union.

As already mentioned, a superordinate social category may be more difficult to visualize, because it is formed by a large number of people, possibly rather heterogeneous. Developing identification with a category of this type first requires developing a perception of its existence. Some studies by Castano *et al.* have investigated how a political institution such as the European Union may be perceived by people not so much as an abstract and far-away entity but as a real community of citizens. These studies have referred to the notion of entitativity, the extent to which a social category is perceived as having a real existence. It is assumed that four factors especially contribute to enhance a social category's entitativity: common fate, similarity, proximity, and boundedness. If we ask people to read a journal article or to see a video stressing either common trends or differences in the economic development of European countries, we are manipulating the common fate factor. If we show people a map in which the boundaries of single European nations are thin lines but the external perimeter of the European Union is a thick line (or vice versa), we are manipulating the boundedness factor. Similarity and proximity may also be manipulated in various ways. Such manipulations have been shown to be effective in increasing people's degree of identification with the European Union. Their effect, however, seems to be felt mainly by people holding moderate attitudes toward the European Union and is not perceived by the so-called Euro-skeptics or Euro-enthusiasts.

Investigation of a higher order identity, such as supranational identity, usually focuses not only on its conditions of existence, but also on its conditions of coexistence with lower order identities. For example, can people strongly identify with Europe and at the same time strongly identify with their nation or region? Studies carried out in different countries have shown that citizens may show high identification with both the superordinate and the subordinate categories. These results suggest that regional, national, and supranational identities may indeed coexist in the citizens minds. This is more likely to happen when the frame of reference is one of integration between different identities, in particular when stressing the superordinate identity does not mean absorbing lower level identities and forcing homogeneity but, on the contrary, accepting a certain degree of internal heterogeneity. Ideally, diversity will become one of the core values of the superordinate identity. This means that belonging to a social category made up of different components that integrate with each other may become a core feature of, for example, North American or European identity. When, however, the frame of reference is one of competition between different identities, in particular when stressing the superordinate identity means absorbing lower order identities or threatening them, competition easily turns into open conflict, and fear of losing a given regional or national identity may hinder development of a supranational identity.

3.3. International Conflict

Analysis of what conditions may favor development of a new supranational identity starting from well-established national identities may be useful in highlighting strategies of international conflict resolution. In his article "International Conflict," Herbert Kelman describes psychosocial processes that underlie conflict insurgence and its escalation, as well as psychosocial processes that may instead favor a break in the spiral of conflict in the direction of conflict resolution. An essential step of an effective strategy of conflict resolution is the acknowledgement of each party's needs. An effective, long-term solution is unlikely to arise if either party has the feeling that its needs or its very existence is being denied or simply ignored by the other party. A similar premise may be found in the dual-identity model of conflict resolution, according to which a first essential step in negotiation between groups consists of creating conditions for the already existing group identities to be expressed and safeguarded. Although open expression of each group's identity and related needs may initially enhance intergroup conflict, it also has the advantage of enhancing each group member's perception of having a well-established identity. This is an ideal psychological premise for an effective interaction with the other party, in order to find solutions that address not only a single party's interests but also superordinate interests. If, however, the first step—expressing one's group identity—is lacking, members of each group may feel that their group's identity is threatened. The resulting uncertainty prevents the group from taking the next crucial step in the negotiation process, finding a solution that satisfies superordinate interests.

4. POLITICAL COMMUNICATION

When we think of politics, we are likely to think of a phenomenon encountered largely through the mass media. The link between politicians, events, and issues on the one hand and most citizens on the other is mediated through the complex web of information generated by newspapers, magazines, television, radio, and the Internet. While the mass media has significantly increased the opportunity for politicians to reach citizens, communicate, inform, and put forward persuasive messages, it has also offered citizens the means to scrutinize politics and politicians more closely while offering information on a previously unthinkable scale. However, if the media multiply the opportunities for communication, it must also be recognized that the process of mediation is not neutral. Through the people who control them, the media select, classify, and shape political information, thus contributing to the creation of political reality.

The influence exerted by the media on political information processing and the degree to which citizens are aware of this influence will be the first issues dealt with in this section. Next, we will consider how political actors communicate and employ language and other communicative strategies in order to fulfill their goals. Our attention will be especially focused on two communicative contexts, political interviews and political discourses.

4.1. Mass Media and Politics

Studies on information processing have shown that a concept or an attitude that has been frequently retrieved in the past becomes more accessible to the individual's mind, that is, more likely to be retrieved again and to be perceived as relevant. Therefore, when media coverage of a given political issue is high, individuals are repeatedly exposed to it, and its accessibility is likely to become higher. Studying the cause–effect relationship between media coverage and accessibility or perceived relevance of a political issue is not an easy task. Is a political issue perceived as relevant because it is widely covered by media, or is it the other way round? Adopting a temporal series technique, some studies have measured perceived relevance of a number of political issues at different moments in time and have monitored media coverage of the same issues in between. Results suggest that media coverage significantly increases the relevance of political issues to which people already attribute a certain degree of relevance. In other words, media have the effect of amplifying the importance of a given issue and possibly of altering the hierarchy of the relevance of political issues. This may have not only generic consequences on the evaluation of political reality, but also more specific effects on the evaluation of a given politician. For example, before the Gulf War, U.S. citizens' evaluation of three main policy areas of President George Bush (foreign affairs, domestic affairs, and economic affairs) equally contributed to form the global judgment of Bush, whereas after the Gulf War, evaluation of Bush was based significantly more on his foreign affairs policy. Thus, media coverage of a political issue leads people to perceive that

issue as more important for the country than they thought before; consequently, that issue weighs more in their evaluation of political leaders.

Are we aware of media influence? Yes and no. We are inclined to think that the mass media influence "people" or "others," but not ourselves. This is the so-called third person effect, according to which we overestimate media influence on others and underestimate media influence on ourselves. This tendency has been defined as a self-serving bias, in the sense that the perception of having full control on the surrounding reality and on one's own choices contributes to raise the individual's self-esteem. Thus, one's capacity of resistance to media influence is enhanced if compared with others' vulnerability to the same influence. The third person effect is more evident in those who perceive the media as untrustworthy and perceive target people as distant from themselves. In a post-electoral survey carried out in Australia by Duck *et al.*, people asked to evaluate the influence of a number of television programs (political debates, news, electoral spots, etc.) on voting choice declared themselves to be the least vulnerable to that influence, followed by voters of their party, voters of the opposing party, and voters in general.

4.2. Political Language

We may undoubtedly say that politics is made up of words. Politics deals not only with concrete, tangible objects, but also with abstract concepts or notions to which people attribute a shared meaning (e.g., democracy, liberalism, welfare state). Moreover, the planning of future events is an essential component of politics, which therefore often refers to something that does not necessarily exist now but that might become real in the future. Such referential elusiveness explains why political language has often been described as polysemous, many-faceted and, in the end, ambiguous. At the same time, such elusiveness makes it clear that political reality is mainly construed through language. Politicians propose a certain interpretation of political issues and events, which may be defined as successful when it becomes shared by a large number of people. Such interpretation often implies stressing similarities or differences between political groups, thus favoring or hindering citizens' identification with these groups.

In studying political language, psychology focuses on the analysis of cognitive, motivational, and psychosocial factors that guide speakers and listeners in the political domain (politicians, activists, journalists, or citizens). Special attention is devoted to the study of the various kinds of communicative contexts in which politicians find themselves. Although each context has peculiar characteristics, politicians' communication is generally guided by three main goals that may be found across contexts: (1) presenting a positive image of themselves and their party, (2) presenting a negative image of opposing parties, and (3) trying to gain consensus. This means that on the one hand, politicians try to define themselves, to make clear what their political views are and how they differ from other politicians' views. On the other hand, they aim to extend the boundaries of their party to include as many people as possible. Language helps politicians in the difficult and contradictory task of defining their group and at the same time trying to widen the boundaries of this group. In the communicative contexts we discuss below, we will consider how politicians defend their image when interviewed and how they use political discourse to define and depict social categories in order to gain consensus.

4.3. Interviewing Politicians

Thanks to the growing media-centered and spectacular character of politics, interviews have become one of the main channels through which politicians present themselves and their political positions. Psychosocial research on question–answer exchanges has shown many, often subtle, strategies employed by interviewer and interviewee to pursue their communicative goals. Extension of this approach to the analysis of political interviews started from two considerations that easily arise when listening to political interviews: (1) journalists' questions are often not neutral and (2) politicians often do not reply to journalists' questions. A systematic examination of interviews given by politicians has led to identifying an interesting typology of questions and answers as well as conditions that make non-replies more likely. As already mentioned, one of the politician's main goals is to present a positive image of him- or herself. According to the so-called face model of political interviews, developed by Bull *et al.*, during interviews politicians tend to defend (1) their personal-political face; (2) their party face; and (3) the face of significant others. Journalists' questions that are perceived as face-threatening are therefore more likely to get a non-reply. Questions may threaten a politician's personal-political face in various ways, for example, by creating a negative impression of the politician (e.g., "Isn't all this emphasis on personality a cover for the

fact that you haven't got a big idea?"), hinting at possible future difficulties, highlighting contradictions with past statements, or stressing difficulties in clarifying personal beliefs (e.g., "I wonder whether wavering voters aren't influenced by not quite knowing where you, Prime Minister, stand"). Similar threats may be found in questions regarding the politician's party (e.g., "Things aren't looking that good for your party, are they?"). Finally, questions may focus on other people or groups but still imply a threat to the politician's face. This is the case for questions implying that the politician does not care for the electorate or support a friendly country, and for questions luring the politician into awkward admissions regarding opposing parties (e.g., "Is there really a shift of opinion toward the Liberal Democrats or is it because they have run a better campaign than you have?", a situation in which confirming either alternative offered by the question would lead the interviewee to support an opposing party).

Although the interviewer has various ways to threaten the politician's face, the politician has various ways to avoid replying. Not replying does not mean being silent, but rather adopting communicative strategies aimed at avoiding a direct reply without necessarily appearing rude or not collaborative. The outcome will be a reply that may be defined as equivocal, because it is not consistent with the interviewer's communicative intention; it is, however, consistent with the politician's intention of presenting a positive face to the electorate. Equivocal replies may include ignoring the question, replying with another question, criticizing the question, or refusing to answer.

4.4. Political Discourse

In their discourses, the politicians' ultimate goal is gaining the widest possible consensus from their audience. They aim to create in their public a political identity that is consistent with their own and that of their party, so as to favor mobilization in their favor. To do so, politicians address not only single individuals but also significant groups to which their audiences belong and are likely to identify with. For example, they can explicitly address blue-collar workers, entrepreneurs, or housewives. They may also address wider social categories, such as the citizens of their country, or some vague social categories, such as all honest people: making these categories salient means addressing millions of potential voters. Whatever social category they are trying to make salient to their audience, politicians will use arguments that highlight the fact that the values and goals of that category are consistent with the ones their party is ready to fight for (e.g., "Citizens want to be able to walk safely in their neighborhoods, and that's exactly what we're aiming for with our new bill on criminal offenses"). Very likely, they will also argue that the same values and goals are inconsistent with those of the opposing party. Whenever this is possible, politicians also try to present themselves as prototypical members of the salient category (e.g., "I am a father/mother myself" when speaking about school reform), thus implicitly stating that they are trustworthy and entitled to represent the category.

A confirmation of the relevance social categories have in politicians' discourses has been offered by some studies that compared how politicians of opposing parties interpret the same political event. Differences were observed regarding not only the description of facts, but also the definition of social categories involved in the event, as well as of the values and goals presented as typical of those categories.

An example may be found in Reicher and Hopkins' 1996 study on how a famous miners' strike in England was reconstructed from opposite perspectives by political leaders Margaret Thatcher (Conservative Party) and Neil Kinnock (Labour Party). In her interpretation of the events, Mrs. Thatcher created a picture in which miners that did not participate in the strike, as well as their relatives, were attributed personality traits such as courage and determination that were presented as typical of authentic English people. Mrs. Thatcher's definition of the category of strikers, on the other hand, identified them with violent people and even with terrorist groups. Thus, through her discourse Thatcher evoked a large ingroup of "no-strikers/authentic English people" on the one hand and a small outgroup of "strikers/terrorists" on the other. She consistently presented the Conservatives' reaction to the strike as a way of defending the nation at large against a minority of destabilizing forces. A similar though opposite argumentative strategy may be found in Mr. Kinnock's interpretation of the same event. He described the strikers as people who wanted to defend their rights against the repressive policy of Mrs. Thatcher. In doing so, Mr. Kinnock compared strikers to all Labour voters and more generally to all the people who were against Mrs. Thatcher, so that the initially small category of strikers ended up being redefined as the large ingroup of "strikers/liberals."

To conclude, the way in which both a political event and the people involved in it are defined is of crucial

relevance to increasing the public's feeling that the politician, the politician's party, and the public belong to the same wide ingroup and share the same goals and values. Politicians try to define the boundaries and the characteristics of social categories in order to mobilize people in a direction that is favorable to them and unfavorable to the political adversaries. Social and political categories are not stable and unchangeable; on the contrary, they are constantly being redefined, and political leaders contribute to their redefinitions through political discourse.

4.5. Political Persuasion

So far, we have focused our attention on how politicians present their political identity and try to shape the public's political identity. Research on political persuasion, however, has been more generally aimed at investigating how politicians' messages may have an impact on the target's (i.e., the public's) political attitudes. The typical research design of this kind of study implies a comparison of the target's attitude toward a given political issue (e.g., the tax system) before and after exposure to a persuasive message regarding that issue (e.g., a political discourse in favor of a reduction of the tax load on wages). Recent studies in the area have focused on the so-called matching effect, according to which the persuasive impact of different types of messages varies as a function of given characteristics of the target. For example, Lavine *et al*. investigated the different persuasive impact that threat-related and reward-related political messages may have on individuals characterized by different degrees of authoritarianism. The messages were intended to persuade young people to vote in presidential elections. While threat-related messages highlighted the possible negative consequences of nonvoting (e.g., "Not voting allows others to take away your right to express your values"), reward-related messages highlighted the possible positive consequences of voting (e.g., "Voting enables one to bring about the kinds of public policies he or she believes in"). The change in attitudes toward voting after exposure to either type of message was then measured, comparing participants with high versus low degrees of authoritarianism. Threat-related messages had a high persuasive impact on authoritarian people and a low impact on nonauthoritarian people, and the opposite was true of reward-related messages. Thus, the matching effect suggests that tailoring political messages with regard to the target's characteristics (e.g., authoritarianism) may help to increase the persuasive effect of those messages.

5. POLITICAL PARTICIPATION

For most people, voting is the only practical way to take part in political life. At the same time, in a democratic regime, it is also the most important one. In this section, we will focus on how voting behavior has been explained, by referring to motivational, cognitive, and social factors that may underlie political choice. Finally, we will briefly hint at more involved forms of political participation.

5.1. Voting Behavior

Since the pioneer work by Campbell *et al*. on party identification, that is, on voting as the stable result of an identification process that dates back to developmental age and to family influences, several other explanations have been proposed. As we have seen in Section 1, the extension of rational choice theory from mathematics and economics to politics has led to the view of voting behavior as the outcome of a rational process. In other words, people consider all the information they have regarding the various possible alternatives (parties or candidates), and choose the one that is characterized by the highest possible utility for them. However, ideas of human rationality and utilitarian goals implied by this approach have been seriously challenged by subsequent psychological and empirical research on voting behavior. The main acquisitions of this research are summarized here.

5.1.1. Cognitive Factors

When deciding how they will vote, people do not necessarily consider all the information they have acquired over time about the various parties or candidates. Given that the quantity of information they can pay attention to at a given moment is limited, they tend to focus only on information they perceive as salient and relevant at that very moment. Their decision-making process does not consider all pros and cons of each alternative. Instead, a series of heuristics or simplification strategies is adopted, which allow making a decision through a limited consumption of mental energy. Thus, people base their decision on a limited amount of information they perceive to be relevant at a given moment. Such information may be categorized into the following four main dimensions:

1. *Trait-based voting*. As we have already seen (see Section 2), candidate evaluation is a very good predictor

of voting behavior. The relevance of this dimension is likely to have increased in recent years, due to the growing personalization of politics favored by the media.

2. *Issue-based voting*. Voting choice may be based on parties' political programs or their positions on given political issues, such as social welfare, the tax system, or immigration rules, and people will choose the party or candidate whose positions on those issues are more consistent with their own. According to the so-called consumer model, recent years have been characterized by an increasing tendency for people to base their voting choice on their agreement with a party position on one or more relevant issues. This would lead to a party choice that is related to specific issues and consequently that is not stable over time.

3. *Ideology-based voting*. In contrast with the consumer model, the ideological model assumes that people choose parties or candidates with whom they share not so much a position on a given issue but general beliefs based on shared ideology or values. As the organization of political attitudes around abstract principles is typical of politically sophisticated people, ideology-based voting is likely to be more concentrated among this kind of people.

4. *Performance-based voting*. Voting choice may also be influenced by the evaluation of past performance of a party or a political leader. Actually, representation of the past is often the main source of information people use to make predictions about the future. Not only reference to the actual past, but also mental simulations of how things might have developed differently, significantly contribute to shape people's interpretation of political reality. For example, one may think that the armed intervention of the United States and their allies in Iraq in 2003 was appropriate if the most readily available alternative to the individual's mind was a course of events implying Saddam Hussein's alliance with Al Qaeda and a rise of terrorist attacks in Western countries. On the other hand, one may evaluate the same intervention as inappropriate if to the individual's mind the available alternative was a successful mediation process handled by the United Nations. A negative evaluation of past action of politicians may sometimes lead to strong disenchantment with politics and increase the possibility of success for new emerging or extremist political parties.

While each of the preceding factors has been shown to play a significant role in people's choices, recent research has been developing more complex predictive models of voting behavior aimed at assessing the relative weight of each factor on voting. Through application of these models, one may be able to assess how much the electoral success of a given party may be attributable to the party leader, to party ideology, to past performance of either the party or the leader, and so on.

5.1.2. Personal Interest

According to rational choice theory, utility pursued in voting decision corresponds to self-interest, intended as material, economic interest. Recent empirical research in political cognition has reconsidered this issue in a different perspective, focusing attention on political attitudes that are perceived by people as personally important. Issues of domestic policy, such as welfare or taxation, are usually perceived as more personally important than other issues. When asked to express their attitudes regarding issues considered personally important, people reply more readily than regarding other issues. This suggests that attitudes toward domestic policy issues are more immediately accessible to the individual's mind and therefore are more likely to influence voting behavior. The same does not hold for attitudes toward issues of foreign policy, such as defense expenses or non-intervention in war. Though these issues are perceived as more important than domestic issues at a national level, related attitudes are less easily accessible to the individual's mind and therefore are less predictive of voting behavior. Meta-analysis of a number of survey data sets in the United States has confirmed that attitudes on policy issues that are personally important are more stable over time and are more predictive of candidate preferences and voting.

Analysis of what factors may be at the origin of a political attitude's personal importance has shown that self-interest, as defined by rational choice theory, is just one of these factors. Another factor that significantly contributes to raise the personal importance of a political attitude is its relation with one or more core values, values that hold a high position in the individual's hierarchy. This means that, for example, some people may perceive having an efficient public health system as important not only because of the personal advantage implied by such a system but also because solidarity is one of their core values.

5.1.3. Social Factors

In addition to self-interest and core values, a third factor that has been shown to underlie the personal

importance of political attitudes is social identification with reference groups. As we have seen in Section 3, social identity is that part of an individual's self-concept that derives from awareness of belonging to one or more groups. If voting for a given party is a shared and relevant characteristic of a group one identifies with, voting for that party may become a way of expressing and reinforcing one's group identity. A confirmation of this can be found in research that has investigated psychological antecedents of the so-called regional vote, that is, the prevalence of a given party in a given region or area of a country. For example, Scottish people are traditionally more in favor of the Labour Party: research results by Abrams and Emler show that Scottish people who vote for the Labour Party have a higher Scottish identity than Scottish people who vote for the conservatives. Although people with higher Scottish identity feel they have fewer economic opportunities as compared to people living in England, these people are also more willing to remain in Scotland. These data suggest that voting may be a way of expressing one's social identity, in this case, regional identity. Even more, they suggest that in voting behavior, expression of group identity may challenge, if not overcome, the pursuit of economic self-interest.

In conclusion, empirical research on voting behavior has shown that vote is not the outcome of fully rational thinking aimed at maximizing personal benefits and minimizing personal costs. It is a more complex process that implies selective information processing, reference to values or other abstract principles, and the expression of one's own identification with one or more social groups. All these factors allows one to overcome the so-called voter's paradox described by rational choice theorists, which portrays voting as an apparently contradictory action: The costs implied by the act of voting (e.g., acquiring knowledge of the issue, renouncing a day's leisure) are higher than the benefits the single citizen may expect from that act, since one single vote cannot decide the outcome of the election. Thus, if voting were based on a mere cost–benefit ratio, no single citizen would vote. As we have seen, however, voting decision is not the outcome of a simple cost–benefit ratio but rather of a range of cognitive and psychosocial factors that must be taken into account if satisfactory predictive models of voting behavior are to be developed.

5.2. Collective Action

When we shift our attention from voting as the basic level of political participation to more involved forms of participation, such as demonstrations, sit-ins, or strikes, finding satisfactory explanations of why people get involved in such actions becomes even more challenging. On the one hand, the cost of similar actions may be very high for the individual, in terms of time, money, and energy spent in participation, but also, in extreme cases, in terms of stigmatization or physical damage suffered because of participation. On the other hand, possible benefits deriving from the success of collective action may be enjoyed just as much by people who did not actively participate in that action, which may be one further reason to not undertake the risk of participation. In his article "Collective Action," Bert Klandermans deals with this issue in detail, focusing on three main motivations that lead people to get involved in collective action: instrumentality, identity, and ideology. Klandermans also offers an articulate definition of collective action, distinguishing it from other types of group actions. Finally, he analyzes a number of factors that play a role in transforming potential participation into actual participation.

See Also the Following Articles

Collective Action ■ Ideological Orientations and Values ■ International Conflict

Further Reading

Abrams, D., & Emler, N. (1992). Self-denial as a paradox of political and regional social identity: Findings from a study of 16- and 18-year olds. *European Journal of Social Psychology*, *22*, 279–295.

Boninger, D. S., Krosnick, J. A., & Berent, M. K. (1995). Origins of attitude importance: Self-interest, social identification, and value relevance. *Journal of Personality & Social Psychology*, *68*, 61–80.

Bull, P., Elliot, J., Palmer, D., & Walker, L. (1996). Why politicians are three-faced: The face model of political interviews. *British Journal of Social Psychology*, *35*, 267–284.

Campbell, A., Converse, P. E., Miller, W. E., & Stokes, D. E. (1960). *The American voter*. New York: Wiley.

Caprara, G. V., Barbaranelli, C., & Zimbardo, P. G. (1999). Personality profiles and political parties. *Political Psychology*, *20*, 175–197.

Castano, E., Yzerbyt, V., & Bourguignon, D. (2003). We are one and I like it: The impact of ingroup entitivity on ingroup identification. *European Journal of Social Psychology*, *33*, 735–754.

Catellani, P., & Quadrio, A. (1991). Ideal and real in the representation of politics. *Revue Internationale de Psychologie Sociale*, *4*, 231–256.

Catellani, P., Milesi, P., & Crescentini, A. (in press). One root, different branches: Identity, injustice, and schism. In B. Klandermans, & N. Mayer (Eds.), *The psychology of right-wing extremism in Europe*. London: Routledge.

Downs, A. (1957). *An economic theory of democracy*. New York: Harper & Row.

Duck, J. M., Hogg, M. A., & Terry, D. J. (1995). Me, us, and them: Political identification and the third-person effect in the 1993 Australian federal election. *European Journal of Social Psychology*, *25*, 195–215.

Gaertner, S. L., Rust, M. C., Dovidio, J. F., Bachman, B. A., & Anastasio, P. A. (1994). The contact hypothesis: The role of a common ingroup identity on reducing intergroup bias. *Small Group Research*, *25*, 224–249.

Granberg, D. (1993). Political perception. In W. J. McGuire, & S. Iyengar (Eds.), *Explorations in political psychology* (pp. 70–112). Durham, NC: Duke University Press.

Haslam, S. A. (2001). *Psychology in organizations. The social identity approach*. London: Sage.

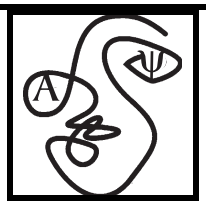
Krosnick, J. A., & Brannon, L. A. (1993). The media and the foundations of presidential support: George Bush and the Persian Gulf conflict. *Journal of Social Issues*, *49*, 167–182.

Lavine, H., Burgess, D., Snyder, M., Transue, J., Sullivan, J. L., Haney, B., & Wagner, S. H. (1999). Threat, authoritarianism, and voting: An investigation of personality and persuasion. *Personality and Social Psychology Bulletin*, *25*, 337–347.

McGraw, K. M. (2000). Contributions of the cognitive approach to political psychology. *Political Psychology*, *21*, 805–832.

Miller, A. H., Wattenberg, M., & Malanchuk, O. (1986). Schematic assessment of presidential candidates. *American Political Science Review*, *80*, 521–540.

Reicher, S., & Hopkins, N. (1996). Self-category constructions in political rhetoric; an analysis of Thatcher's and Kinnock's speeches concerning the British miners' strike (1984–5). *European Journal of Social Psychology*, *26*, 353–371.



Postmodernism in Psychology

Teresa Cabruja Ubach
University of Girona, Girona, Spain

1. Introduction
2. Contextualization: Issues of Postmodern Thought as a Critique of Modern Rationality
3. Postmodernism Comes to Psychology: The Crisis of Modern Psychology and Its Different Theoretical and Practical Proposals
4. Implications for Traditional Psychological Topics Such as Emotion, Attitudes, and Memory
5. Implications for Psychological Fields and the New Role of the Psychologist
6. An Example: Implications of the Postmodern Turn on the Study and Conception of Identity and Gender
Further Reading

GLOSSARY

critical psychology Psychological theories that try to substitute the study of psychological processes used in positivist psychology for a study based on language and social practices.

cyberpsychology An approach that allows thinking about technoscience, cultural theory, and political effects.

deconstruction Action to disassemble structures that are supposed to be well known in order to know the non-explicit way in which they have been organized and the premises that construct them.

discourse analysis Variety of forms developed by linguistics, philosophy, and psychology to analyze both social linguistical practices that produce social meaning and the effects of the discourse.

enlightenment Characteristics of psychological knowledge, producing changes on studied objects that have created this knowledge.

postmodernism Movement or set of ideas that implies an epistemological and social change with respect to the basic principles of the modern rationality.

social construction Idea that human and nonhuman objects do not exist independently of human actions and their linguistical practices.

Postmodernism can be understood as the critique of modern scientific rationality. It may be thought as a loss of belief in truth, in universalistic ideas of historical progress, in unitary subjects, and in objective scientific method. This also means incredulity regarding the great narratives of legitimacy. Postmodern thought contains diverse interpretations and proposals that replace a positivist method of knowing the "real world," by local and engaged accounts of reality with focus on social and linguistic construction. The influence of these ideas can be found in psychology, as in any other discipline. Here, it reflects, among others things, shifting conceptions in the main epistemological and methodological topics that comprise the field of psychology. Postmodernism can be found overall in the social constructionist approach (Gergen, Ibáñez, etc.), but also in others' approaches. This article introduces the general themes in postmodern thought, and its bases in French poststructuralist philosophy, feminist studies, and postcolonial studies. All these have had their effects on psychology and we shall see

©2004 Elsevier Inc.
All rights reserved.

how its principal topics of study such as emotion, attitudes, memory, self, and gender have all been challenged.

1. INTRODUCTION

During the 1970s, a "crisis" in psychology reflected a more general crisis in the social sciences: the crisis of the paradigm of positivist modern rationality as a privileged form of knowledge of the human being. This debate was inspired by French poststructuralist philosophy (Foucault, Derrida, etc.), feminist studies (Kristeva, De Lauretis, etc.), and postcolonial studies (Gayatri-Spivak, Bhabha, etc.). The postmodern thought was represented by Lyotard, Baudrillard, Haraway, Bauman, and Butler, among other thinkers.

This crisis emerged from the recognition that a general openness to change had, by the 1970s, precipitated transformations in both the academic and the personal spheres under the influence of social movements such as feminism, ethnic movements, and the political, social, and cultural transformations that drove the new technologies.

Events of the 20th century have prompted several questions about truth, science, reason, and the concept of reality. Within the social sciences, a critique has been developed, based on the questioning of the basic dichotomies of Western modern thought such as reality–fiction, nature–culture, emotion–reason, objective–subjective, and mind–world.

Mainstream psychology, until the end of the1970s and the beginning of the 1980s, had followed the dominant paradigm of the natural sciences. It put its faith in the hypothetical-deductive method, a methodology for human research that separates science, on the one hand, from "the world" on the other. The fundamental ideas that constitute this form of knowledge are objectivity and neutrality. These are supposed to produce knowledge that is both universal and predictive. Nevertheless, in different fields of psychology, doubts have arisen as to whether objectivity and neutrality can handle complex reality. These reframings suggest that we see psychological knowledge as impregnated of values and "experimental emptiness." In psychology, this is signaled, for example, in the emphasis on the Western cultural, gender, and sexual values as individualism, ethnocentrism, logocentrism, and androcentrism that are inherent to psychological practice and research. Psychology is challenged in its claims to objectivity and neutrality and in its belief that the world and the subject are independent from each other.

The aim of change can be seen clearly in the works of Gergen, in his treatment of psychology as history, in the socioconstructionist paradigm, and in the analytic critiques of experimental psychology. Challenges to traditional psychology also exist in Latin America with the development of a more reflexive and engaged communitarian psychology. It is important to emphasize that of the several ways aimed at changing the traditional psychological form of thinking and doing research, only some are directly acknowledged as postmodern. Other are recognized as "critical psychology" (among other terminologies).

General topics that are often taken for granted in psychology have begun nowadays to be deconstructed. In other words, this postmodern turn shows how the system of representation of the world is not free of values and can operate by promoting or oppressing certain forms of life or personhood, within the hierarchy of power relationships.

2. CONTEXTUALIZATION: ISSUES OF POSTMODERN THOUGHT AS A CRITIQUE OF MODERN RATIONALITY

2.1. Postmodernity and the "Postmodern Condition"

Postmodernity is a controversial term. The French philosopher Lyotard claimed that it neither defines nor identifies a period or a historical time. Postmodern rather serves to indicate that something is in decay in modernity, in the universality of the philosophical discourse of the enlightenment. Terms and their opposites begin to lose their sense and enter the play of autoreflexivity, namely, the differences between fact and fiction, language and its objects, art and "kitsch," as well as the limits between styles, disciplines, genres, etc.

How can postmodernity be understood? Postmodernity can be found in the goal, common to literary theory, philosophy, and the social sciences, to reveal how it is that positivist dominant discourses are able to create "illusions" of theoretical coherence and empirical credibility from certain institutional conditions and with consequent effects of power.

The post-structuralist, feminist, and post-colonial contributions constitute a source of inspiration for the "postmodern turn" as well as the new sociology of science and neopragmatism. The philosophical base of the modernity–postmodernity debate is found in

French poststructuralist thought in its various forms. In particular, the work of Foucault on matters such as power, madness, sexuality, prisons, and subjectivity develops the thesis of the relationship between power-knowledge-truth and the political, historical, and rhetorical character of psychological sciences.

One of the key methods of postmodernity is the deconstruction of Derrida and one of his main ideas: that language does not have any sense or truth, directly or literal, and its relation with social institutions and hierarchical and power relations. A continuous self-reflexivity is proposed. One consequence is to consider scientific knowledge as a form of knowing the world that is not beyond any other human activity.

What is the postmodern condition? According to Lyotard, it is understood as the generalized acknowledgement of the exhaustion of reason and is based in the crisis of "representation." That is, it includes the crisis of the social and political movements and discourses, as they are constructed from modern rationality, as well as the transformations in social relations, changes of production, and the global culture of new technologies.

2.2. Three Main Ideas from Postmodern Thought

To be precise, the postmodern critique of modern scientific rationality involves the following three ideas.

1. The critique of great narratives that legitimize a multitude of social actions, erected as discourses of "truth," and of the linear and evolutionary ideas of progress and history inherent to the modern project. Instead, a local and committed knowledge is proposed.
2. The critique of the concept of representation that entails the dissolution of the border between the language and its object and of the idea of mind as the location of knowledge. A socially constructed reality is proposed.
3. The critique of the subject of Western knowledge and the vindication of "differences," and the death of the subject as the center of representation and history, and the recognition of its universalism and androcentrism. The consequences for emancipatory social movements appear controversial. The proposal is to avoid essentialism and universalism and to consider using identity categories of a more temporary and local dimension.

With these three critiques, statements of "objective truths" become constructions of communities with particular values; interest moves towards the forms in which the world "is represented" and what consequences these representations carry. This will bring about a change in the traditional assumption of the way in which language operates. It will no longer be seen as an instrument for the expression of internal representation of external reality, but rather something that obtains its meaning from its use in social life. What is in play, then, is the notion of objectivity as a legitimacy of the truth about how the world and what people are like.

3. POSTMODERNISM COMES TO PSYCHOLOGY: THE CRISIS OF MODERN PSYCHOLOGY AND ITS DIFFERENT THEORETICAL AND PRACTICAL PROPOSALS

3.1. Modern Psychology and the Analysis of Its Own Discourse

With the postmodern turn, psychology, from the reflexivity and the deconstruction point of view, analyzes its own discourse. It tries to understand the way in which it opts to produce certain knowledge and what relations of power are produced. Social constructionism in psychology requires that we put aside the basic presupposition of modern rationality that, as objectivity, had led psychology to constitute itself, as Tomás Ibáñez indicates, as an extraordinarily authoritarian system. Psychological entities were naturalized as corollaries of the human condition, pressuring it to adapt itself to them and freeing the psychologist from the responsibility of the effects of power over the individual. Thus, the productive character of psychology has made a vocabulary and new knowledge of the individuals possible. But simultaneously, its commitment to a rhetoric of truth has entailed effects of social regulation. An individual's motivations or attitudes, satisfaction, self-esteem, etc., are the elements that are used in deciding who gets a job, who may look after their children, etc.

To consider, then, that psychology is not an objective science like the natural sciences entails a series of changes. Is the conduct of human beings really predictable and not altered by people's knowledge of themselves, or by the continuous transformations of the world around them? Kenneth Gergen calls this effect enlightenment.

3.2. The New Conception of Psychology: The Postmodern Psychology

Due to its symbolic character, situated in a concrete society, culture, and history, psychology's theories, from a postmodern point of view, are to be understood as a practice, and its concepts are to be understood in a dialogical form. For example, if a complex, a frustration, a motivation, an attitude, or a phobia exists, it is because psychology segments, differentiates, and compartmentalizes certain phenomena and gives them a name. It divides populations and their actions into normal and pathological, or normal and abnormal. The possibility of decoding an interview or an observation of a child in a daycare center depends on training that has separated out different behaviors, compared them, named them, and classified them. In other words, it has made available to the people trained in this discipline the authority to carry out the actions that correspond to this domain.

An important change is to stop conceiving the concepts developed by psychology as individualized and as the result of cognitive or cerebral operations. Rather, they should be thought of as relational and evaluative, or in other words, as ways to understand the world and to agree or disagree with it, or to act on it or cause reactions in it. In addition, this way of thinking about psychology totally changes the form of acting and producing personal and social transformations. In the field of psychotherapy, as in educational or social psychology, it leads to specification of from where and why the psychologist wants to intervene, not from an aseptic and neutral form, but rather from an engaged and participative position.

Postmodern psychology replaces a concern for the subject of investigation and the adequacy of theories with an interest in the procedures and instruments used for the study of the socially constructed subject and the way in which we relate to it. And, above all, it is interested in the social conditions that give rise to such a subject and its relations. The concepts inherited from positivist psychology are reanalyzed, and the "great narrative" of psychology is replaced with local knowledge. The attempt to find universal laws is dropped in favor of circumscribing knowledge to certain socio-historical conditions of production, susceptible to change. It has been suggested that the professional practice of psychology may get along well with postmodern thought, since it focuses on "situated" knowledge of human activity, the everyday world, and validation through practice; dynamic, dependent on context, and fragmentary.

4. IMPLICATIONS FOR TRADITIONAL PSYCHOLOGICAL TOPICS SUCH AS EMOTION, ATTITUDES, AND MEMORY

From postmodern psychology's point of view, our conception of self, science, representation, and language influences the way that the human being has been studied. This implies a deconstruction of the typical themes of the discipline. Identity, emotions, gender, attitudes, and memory are some of the subjects that have been reconsidered along the lines previously mentioned.

4.1. Bases for Reconsidering the Main Topics of Psychology

1. Proposing a less individualistic and more collective or relational conception;
2. introducing a dialogical perspective, a more contextual and a non-universalistic or essentialist analysis;
3. being conscious of one's own prejudices of interpretation;
4. recovering the role of language not as a simple transmitter but as constructivist;
5. introducing the creative, resistant, and transforming aspects of subjects; and
6. introducing a dimension that defines the moral, evaluative, political, and functional character that must lie in any object of analysis within a certain social order and intersubjective relation.

4.2. Examples: The Study of Emotions, Attitudes, and Memory

In the study of emotions, they are not seen as responses or reactions, but as a part of a contextual representation, not in the intentional sense, but rather as significant in each different situation. In this sense, the emotions are considered as having performativity and pragmatic meaning, showing how the individual understands their context and how they express moral valuations.

Another example is the study of attitudes, almost always considered to be prior to performance, and locatable. But they become understandable as dilemmatic and challengeable constructions, dialogical in character, and as possible *a posteriori* elaborations of our behavior. This way, when we try to understand an event, we relate it to others or to ourselves, resorting to

this cause–effect relation in which we have been socialized. Thus, perhaps we looked for motives and reasons that were not always there before, and which are perhaps reconstructed in some cases *a posteriori*, and that are more applicable to current circumstances than an expression of a previous reason or idea.

In the study of memory, individual memory is not considered as an individual cognitive property, but rather in relation to its function and its use in a certain context. In the production of a particular memory, what is more important than its activation or configuration is the moment and the other people involved. This way, the details, the precision, or the changes in narration are not established by the ability of memory to reproduce reliable truth, but in its relative performativity at the time of production. That is to say, it serves to say something about the individual or to obtain something of the others, and obeys a system of expression of values, beliefs, and functions. Something similar happens in the case of collective memory that is considered in relation to its historic and political role in a concrete socio-economical and political context, with a sense only within a political, economic, institutional and communicative relationship. The implications, for example, in individual or family therapy's ability to solve conflict or trauma are clear, as are the implications in situations of intergroup conflicts and in institutions.

5. IMPLICATIONS FOR PSYCHOLOGICAL FIELDS AND THE NEW ROLE OF THE PSYCHOLOGIST

5.1. Implications of the Postmodern Turn for Psychological Fields

What does this postmodern movement imply for treatment and therapy, for educational, environmental, psycho-social, and organizational intervention, or for the resolution of problems, mediation, and so on? That is to say, what does it imply for applications other than psychology? Since, as we have commented, from the postmodern psychology point of view, any type of individual or group psychological intervention is not considered merely an "application" of the theory, the psychologist cannot be placed in an external relation—as could happen with other sciences—and the values that guide an application and the directions of the transformation need to be specified.

As examples of methods, even if they do not have their origins in postmodern thought but are developed or used by postmodern psychology, we have discourse analysis, with a variety of schools and developments, which is centered on identifying the discursive practices in relation to their functions and pragmatic effects, and conversational analysis, with its emphasis on the interpretation and meaning of the participants in a concrete context and in its institutional background. Both constitute bases for the study of daily interactions and institutional contexts in research, therapy, education, racism, etc., to transform concepts and methods of work into the day-to-day practice of the psychologist.

An example is the communitarian psychology in Latin America, based on the Action-Research of Lewin and ethnographies that have developed the theology of liberation or the socioconstructionist movement. Its basic characteristic is to obtain a more dialogical relation of a horizontal nature. Another example is in psychotherapy, with the reflective teams, inspired by Andersen and White, in which interdisciplinary teams of psychologists and relatives use these developments to jointly find formulas of dissolution or changes of problematic aspects or in work on the psychologist-client dialogue developed by socioconstructionist therapy. In family therapy, counseling, and individual psychotherapy, the emphasis moves away from the individualized problem and is related to cultural and institutional practices, opening therapeutic work to social justice and to larger sociopolitical and relational contexts. Works is also done on what is defined as "problematic," to reconstruct it from alternative forms or, even, deconstructing diagnoses to make possible different solutions.

5.2. Implications for the New Role of the Psychologist

To summarize, some of the proposals of postmodern psychology to equip psychology with a less authoritarian and more democratic, plural, and participative position are related to the new role of the psychologist.

1. The role of the psychologist is understood to be that of an agent who facilitates personal transformation and social change, while being aware of the values and interests of professional performance.

2. The activity of the psychologist depends on the participation of the people involved in the investigation or the intervention, co-producing the obtained knowledge and not holding a privileged position as expert. The procedure should include, for example, the people involved in the decision making and the

evaluations before, during, and after treatments and interventions.

3. The psychologist should be sensitive to diversity and differences as well as the plurality of opinions, situations, etc.

4. A role is developed to commit with its ideas by its professional action, which makes explicit and does not hide under neutrality or aloofness.

5. The psychologist develops a less authoritarian intervention in respect to the status that psychology has in our society, beyond the pursuit of the deontological codes of the professional practice.

6. The psychologist incorporates the personal or communitarian explanation of what happens.

7. The psychologist should incorporate a constant reflexivity in all actions and professional practices in order to question and to query the methods of investigation or intervention.

8. The psychologist should develop and apply methods and techniques that incorporate parts of this program.

6. AN EXAMPLE: IMPLICATIONS OF THE POSTMODERN TURN ON THE STUDY AND CONCEPTION OF IDENTITY AND GENDER

6.1. The Deconstruction of the Subject in the Postmodern Perspective and the Changes with Regard to Identity

In fact, this critique of psychology has been made by a great number of psychologists influenced by symbolic interactionism, ethnomethodology, poststructuralism, and deconstruction, who have questioned the individual/society dualism, the idea of the mature, autonomous, and independent person who is seen as an agent planning his or her life, and who has values in virtue of which he or she makes choices (the self-sufficient individual). Consequently, the idea of evolution or development of modern rationality that impregnates a great part of evolutionary theories of personality or child development is replaced by more relational options and different understandings. The works of Foucault on the technologies of the self, developed by Rose, Walkerdine, Sampson, etc., show that the knowledge of our individuality (for example the search for "our own identity" or the wished-for social objectives, such as well-being, happiness, efficiency, etc.) has been constructed in relation with the use of the capacities of citizens and also for its regulation. The idea of a "saturated self," as Gergen proposed, as the result of a social saturation process, with the new technologies and the variety of human relationships, leads to a condition in which doubts are instilled about who and what we are. This is not the result of our personal essence: real feelings, deep beliefs, etc., but of how we are constructed.

In particular, these developments imply that we should consider a more relational and dialogical "self" and the replacement of individualized theories of identity with more relational theories. A consequence is a work with a symbolic plurality of meanings with its respective positions in networks of power, shared or constructed between different people, present or not. When constituting intersubjectivity in the interaction, "self" can be understood discursively as a "text," constructed by a multiplicity of narratives, about how we relate ourselves, full of other stories that we receive explicitly and implicitly throughout our life, among which it is possible to register not only those of human actors, but also those of nonhuman (nature, for example), as proposed by "actor-network theory." The metaphor of the text is used to talk about the identities in a continuous process, in an intimate relationship with the social institutions and practices. One works with the change from applying different techniques to narrate in another form the events, and thus to transform them.

6.2. Cyberpsychology and the Deconstruction of Gender

The idea of biotechnological self, grounded in the cyborg, an organism that the anthropologist Donna Haraway developed in her study on primates, is an organism made up of technology, biology, culture, and politics, and is used to destabilize the essential differences between body and mind, body and machine, and masculine and feminine. Also, it brings closer another form of new identities that start from, for example, the use of prostheses. Additionally, we can find changes for the gender relations as in cyberpsychology. Cyberpsychology allows us to work with technology and subjectivity, and in politics, feminism, and technoscience with transgressive identities. In fact, a postmodern feminism, as psychologists Raquel Hare-Mustin and Jeanne Marecek suggest, is brought about through a deconstruction of the polarity of sex and

gender, to go beyond the different previous totalitarian theories while maintaining a feminist critique on the effects of how psychology has constructed sexual and gender difference.

See Also the Following Articles

Attitude Measurement ■ Attitudes ■ Cyberpsychology ■ Emotion ■ Gender and Culture ■ History of Applied Psychology, Overview

Further Reading

Gergen, K. J. (2000). *An invitation to social construction*. London: Sage.

Gergen, K. J. (1994). *Realities and relationships: Soundings in social construction*. Cambridge: Harvard University Press.

Gergen, K. J. (1991). *The saturated self: Dilemmas of identity in contemporary life*. New York, Harper Collins.

Gordo-López, A., & Linaza, J. (1996). *Psicologías, discursos y poder*. Madrid: Visor.

Gordo-López, A., & Parker, I. (1999). *Cyberpsychology*. London: Sage.

Hare-Mustin, R., & Marecek, J. (1990). *Making a difference. Psychology and the construction of gender*. London: Yale University Press.

Henriques, J., Hollway, W., Walkerdine, V., & Venn, C. (1984). *Changing the subject: Psychology, social regulation and subjectivity*. London: Methuen.

Ibáñez, T., & Iñíguez, L. (1997). *Critical social psychology*. London: Sage.

Ibáñez, T. (1994). *Psicología social construccionista*. Guadalajara: Universidad de Guadalajara.

Kvale, S. (1992). *Psychology and postmodernism*. London: Sage.

McNamee, S., & Gergen, K. J. (Eds.) (1992). *Therapy as social construction*. London: Sage.

Middleton, D., & Edwards, D. (1990). *Collective remembering*. London: Sage.

Montero, M. (1994). *Construcción y crítica de la psicología social*. Barcelona: Anthropos.

Nicholson, L. J. (Ed.) (1990). *Feminism/Postmodernism*. New York: Routledge.

Parker, I. (1999). *Deconstructing psychotherapy*. London: Sage.

Potter, J., & Wetherell, M. (1987). *Discourse and social psychology*. Sage: London.

Rose, N. (1989). *Governing the soul. The shaping of private self*. London: Routledge.

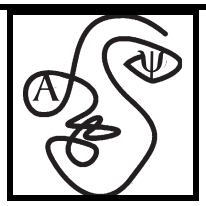
Shotter, J., & Gergen, K. J. (Eds.) (1989). *Texts of identity*. London: Sage.

Shotter, J., & Parker, I. (1990). *Deconstructing social psychology*. London: Routledge.

Simons, H. W., & Billig, M. (Eds.) (1994). *After postmodernism*. London: Sage.

Walkerdine, V. (Ed.) (2002). *Challenging subjects: Critical psychology for a new millennium*. New York: Palgrave.

Wetherell, M., Taylor, S., & Yates, S. J. (2001). *Discourse theory and practice: A reader*. London: Sage.



Postsecondary Education Students, Counseling of

Kathleen L. Davis and P. Gary Klukken
University of Tennessee, Knoxville, Tennessee, USA

1. Overview: The Need for Counseling
2. Problems That Interfere with College Goals
3. Counseling Resources
Further Reading

GLOSSARY

adjustment Psychological processes and behaviors that people use to manage or cope with the demands and stressors of everyday life.

anxiety Feelings of excessive apprehension and worry often accompanied by physical disturbances such as racing heart and sweating.

attention deficit/hyperactivity disorder (AD/HD) A mental disorder characterized by difficulties in maintaining attention with or without hyperactive behaviors.

bipolar disorder A mental and physical disorder characterized by very low mood with intermittent high mood swings; it was previously called manic–depressive disorder.

depression A mental and physical disorder characterized by mood disturbances such as persistent feelings of sadness and despair with a lack of energy.

Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) The classification manual used by health professionals to identify people's mental disorders.

stress Environmental demands made on individuals that may exceed or tax their ability to adapt, cope, or adjust.

Students engaged in postsecondary education often are confronted with stressful situations. Although many students adjust well to the college situation and have the coping mechanisms in their repertoires to successfully negotiate/navigate the stressful events on their own, some students do not. Postsecondary institutions recognize the potential for students' problems that may interfere with institutions' and students' educational endeavors by offering assistance, primarily in the form of counseling. This article reviews normal development and stressors usually encountered by college-age or older adults. It also recognizes many problems with which students present when seeking counseling.

1. OVERVIEW: THE NEED FOR COUNSELING

The majority of high school students planning to continue their education tend to enter a university, community college, junior college, or technical school during the fall immediately following their high school graduation. However, there is also a growing population of high school graduates who enter a postsecondary institution after being employed for a number of years. This article focuses on students seeking postsecondary education at a university or 4-year college, the

©2004 Elsevier Inc.
All rights reserved.

problems that interfere with students reaching their goals, and counseling services offered on campus.

1.1. The Role of Stress and Adjustment

Students who begin immediately after high school graduation and those who resume their studies after a hiatus from formal education naturally will encounter various sources of stress. Stress may be experienced in relation to external events (positive and/or negative) or to internal psychological and developmental issues. During the late 1960s, Holmes and Rahe developed a list of traumatic life events that tend to cause stress in most people. The list of events included things such as changes in schools, living conditions, social groups, and work hours or conditions. Many of these events occur when students begin or are engaged in their postsecondary education. Students who move away from home or lose their support groups in other ways must adjust daily to new situations, including different people (e.g., roommates, various campus groups). Some must split their time between work and school. All experience pressures to succeed in the academic environment. These pressures may be self-imposed or a result of parental expectations. Additional pressure to maintain high grades is placed on scholarship recipients. Many students must reevaluate their view of self in regard to their strengths and skills. Individuals who were identified as extremely successful, academically, politically, and/or athletically, might now be among hundreds of students with similar skills, achievements, and abilities. For someone who has been successful in the world of work and "the one in charge," there now is a need to make conceptual adjustments to being in the student role that often requires more passive behaviors. These changes in environment often induce stress in many individuals. Therefore, in response, students must make adjustments and learn new coping behaviors.

Developmental theorists, such as Erikson and Marcia, suggest that identity concerns are very prominent and normal during late adolescence (i.e., college age). This period is one of questioning and exploring personal values. Values may be examined within the context of areas such as religion, gender, sexual orientation, and potential careers. In addition, as college students enter early adulthood, the quest for intimate relationships becomes paramount. The search for identity and intimate relationships is potentially stressful and anxiety producing, but it is part of the normal developmental process.

Although these stressors are to be expected, individuals often need assistance in learning to adjust or cope with the external and psychological stressors associated with college. There are many ways in which people respond to stressors. How individuals adjust or fail to cope with stressors is dependent on how they appraise and what meaning they attach to the demands and pressures associated with the potential stressors. The perceived degree of threat to one's well-being and coping abilities is related to each individual's previous successful or unsuccessful coping behaviors and experiences. For example, two people may assess attendance at a social event differently. One may be excited and look forward to a wonderful evening, whereas the other may view an evening interacting with others with trepidation. In general, individuals tend to respond to stressful events by developing preferred coping methods or various degrees of maladaptive reactions. It is important that students recognize when they can benefit from assistance in dealing with academic, relationship, and/or personal stressors. Students may seek counseling as a means of learning new ways of dealing with new situations or in recognition of their own inability to adjust effectively to this developmental period and the college experience. Counseling can be preventive, developmental, or remedial.

Prevention or outreach activities of counseling psychologists may entail services such as making a presentation to a dormitory, fraternity, or sorority on the identification and avoidance of date rape or eating disorders. Examples of developmental activities include offering support groups on improving social skills and on learning methods for time management. Training others (e.g., resident hall personnel) on communication skills and peer counseling may be preventive as well as developmental. When a student or others close to the student believe that a problem exists, remediation is usually the typical intervention.

1.2. Influence of Diversity

In addition to the stress and adjustments with which college students must deal, students also bring a diversity of experiences and backgrounds to this new setting. Increasingly large numbers of students are not in the traditional 18- to 22-year-old age range; are from diverse ethnic, racial, and cultural backgrounds; and bring backgrounds of widely varying experiences to the college classroom. Their family systems may range from close and supportive to demanding and

abusive. Their cultural backgrounds may or may not be supportive of higher education. They may have lived on their own or never left home. They may have worked in significant jobs or had careers before going back to school. These students bring new expectations, new goals, and new issues. As a result, campuses are no longer islands of simple intellectual pursuit and career development (if they ever really were); instead, they are dynamic and demanding social structures that must be negotiated on a daily basis.

2. PROBLEMS THAT INTERFERE WITH COLLEGE GOALS

The issues of stress and adaptation appear to have been increasing in intensity over the past decade or so and have been the subject of considerable discussion among higher education mental health professionals. This section briefly examines some of the major problem areas resulting from this increase.

2.1. Academic Problems

Academic problems are common among students. This is the area involving the job of students, that is, to learn new information, new concepts, and new skills. This is also the area for which there is constant evaluation of success or failure. Most students go on to postsecondary education because they did well in prior academic settings. However, being successful in secondary education can happen for a number of reasons, not all of which are useful at the next level. Being academically successful due to easy classes, special privileges, extracurricular renown, or any other basis that resulted in background deficits is a major detriment to college success. Following are some of the most common academic problems and the application of psychology to their resolution.

2.1.1. Study Skills

Many students have not learned how to study. They are usually very intelligent and have used their ability to memorize quickly and/or to write convincingly to succeed in high school. When faced with increasingly difficult material at a college level and competing against equally intelligent classmates, they discover that the tasks of taking good notes, writing college-level papers, and studying for difficult exams is very different from what they were used to doing in high school. Classes and workshops in study skills are offered at most schools using information and techniques developed as a result of psychological studies in human learning. Some of these include overlearning, multichannel information acquisition, and note-taking organization.

2.1.2. Time Management

Many people are very poor time managers, and students are no exception. Choosing what to do and when to do it is a constant decision-making process and is a skill that many students do not have. Students begin to feel overwhelmed and out of control of their lives as academic demands increase in the context of everything else in their lives. Many schools offer training in time management using psychological techniques for values clarification, task organization, and assertiveness training.

2.1.3. Student Role Conflicts

As mentioned previously, the role of students is to learn. This role can be especially difficult for returning students who may have spent years working in jobs, directing other people, and/or being experts in their fields. Although the return to school may be the fulfillment of their dreams, the return to the role of students may be much more difficult. They quickly discover that students are not in charge of the curriculum, do not get to choose their professors' teaching styles, are not experts in every field, and do not get to decide the merits of their own writing styles. In addition, they learn quickly that the way in which they learned to do something on their jobs might not be acceptable in their academic programs. This readjustment of roles may require individual or group counseling focusing on issues of self-identity, motivation, and role flexibility.

2.2. Career Decision Problems

For much of human history, there was not a lot of choice in careers. Usually, an individual's occupation was determined by what his or her parents did as an occupation. Although this process did not provide much freedom of choice, it also eliminated the stress of decision making. When a different occupation was chosen, it was usually based on community need. That is no longer the case in our current society. Each year, there are new

possibilities for careers and old careers may disappear. The result is that decision making about one's career has become difficult and stressful. Following are two types of career decision-making problems.

2.2.1. Undecided Students

There are students who come to school knowing what they want to do for careers and who then continue on into those careers. These are a minority of all students. Many students come to school with no idea of what they eventually want as careers. Others come with ideas or even clear goals and then change their minds. Most students change their intended majors several times over their college careers. Some of these decisions are a result of developmental changes that redirect their interest, whereas others are pragmatic decisions based on classes that they discover they really enjoy, do not like, or cannot pass. There are a number of career development theories about how people make these important life decisions, and counselors will often spend time with students using psychological tests to assess their strengths and interests, provide them with information about employment opportunities and forecasts, and help them to evaluate their past successes as indicators of future possibilities.

2.2.2. Indecisive Students

There are many students whose problems with career choices stem from difficulties with decision making of any kind. Because their career choices are such significant decisions, their indecisiveness may completely stop their progress, eliminate their motivation, increase stress, and cause them to change their majors whenever something new attracts their attention. Counselors are able to make use of knowledge about the fear of failure, the fear of disappointing loved ones, and other unreasonable expectations that may stop students from making decisions. The principles of cognitive–behavioral psychology have very direct application to this problem.

2.3. Personal Problems

All individuals bring their own histories, their own traits, and their own problems with them wherever they go. College is a place where stress can amplify these issues to such a degree that they can interfere with life, sometimes in a serious way. Following are some of the typical problem areas and the application of psychology in their amelioration.

2.3.1. Relationship Problems

Problems with relationships are perennially one of the major concerns for college students. Most students are involved in, or at least concerned about, their search for life partners. Consequently, social relationships are a significant part of their world. They may be dealing with roommates for the first time in their lives and working out the day-to-day adjustments of living with other people who are not family members. They may be in intimate relationships where they must make decisions, set boundaries, and take risks all on their own for the first time. They often hold memberships in various groups such as classroom teams, social organizations, and athletic teams. Here, they must negotiate the more complex social politics of group dynamics. There are also relationships with authority figures that are unmediated by parents or other hometown supporters. Professors have a wide range of personal styles, pedagogical practices, and worldviews. Residence hall personnel bring their own customs and communication styles. Offices of deans and other college administrators have their own agendas and rules. Through this often confusing maze of people, students must successfully negotiate a variety of relationships. It is not surprising that they run into difficulties. This is an area where it is critical to have counselors who are well versed in the application of group dynamics and interpersonal relationships.

2.3.2. Family Problems

Most students are still a part of family units. These special relationships are very powerful, and students' advancement to postsecondary education does not guarantee that these relationships will be trouble free. Some of these family relationships are problematic from the beginning. Many students come to college with histories of abuse. These childhood experiences usually result in long-term damage to self-esteem, interpersonal relationships, and general psychological health. Students with abuse histories may discover their learned wariness of others creates relationship difficulties in many areas of their lives. These problems often require extensive individual or group therapy to explore and work out new patterns of thinking, feeling, and behaving.

Family relationships that have not been problematic might still create considerable stress for students.

Parents do not receive training in how to let go of the more directive role that they have played in their children's lives for many years; however, abandonment of all control is not a healthy solution either. Thus, a new version of old relationships must be developed, often without the mutual agreement of both parties. This is very stressful for both parents and students. It is a developmental issue that everyone must face. Students want more freedom than they had previously (especially those students who are now used to living on their own), whereas parents are often concerned about, or even frightened by, the lack of experience or appropriate behavior exhibited by their sons and daughters. Issues of independence, control, acceptance in a new role, and responsibility for their own lives are very important to resolve as students mature and take their place in adult society. Counseling must be readily available for short-term interventions, with attentive listening mixed with appropriate support, confrontation, and reality testing. Most students survive this stage without the help of professionals, but some outside help at this point can often make the passage smoother and quicker.

2.3.3. Alcohol Use and Abuse

The use of alcohol among students in postsecondary education settings is both a time-honored tradition and a serious mental health problem. Many students use this freedom from parental oversight to increase their drinking behavior. Usually, students do not seek help for problems in this area of their own volition but rather are forced into counseling by concerned others. This may be problematic behavior for only a small percentage of people, but it is a problem that may cost them their academic careers and considerable legal and relationship problems. Although the application of behavioral techniques may be useful here, often more intensive inpatient rehabilitation is needed.

2.3.4. Depression

Perhaps the most prevalent issue in any higher education counseling service is depression. During recent years, this has even exceeded the traditional number one issue of relationship problems. Depression is identified by the fourth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV) as having a number of the following symptoms: depressed mood much of the time, diminished interest or pleasure in nearly all activities most of the time, significant change in appetite, significant change in weight, insomnia or hypersomnia, psychomotor agitation or retardation, fatigue, feelings of worthlessness or excessive guilt, diminished ability to think or concentrate, and recurrent thoughts of death or suicide.

Although there is considerable discussion in the professional community about the prevalence of biochemical depression, it is clear that depression may have either or both a psychological basis and a biochemical basis. The evaluation of a student with symptoms of depression will often include an exploration of current and past stressors, a personal and family history of depression or other psychological issues such as alcoholism, and the current or past use of medication. Depression may have a range of impacts on college students, from simple lack of energy to complete withdrawal from interaction with anyone and serious contemplation of suicide. This can result in academic dysfunction, with concomitant course failures, academic dismissal, and financial repercussions. In addition, interpersonal relationships may deteriorate, social and organizational obligations may be ignored, and physical health may be compromised. Suicide is the most dangerous outcome of depressive behavior; therefore, considerable attention must be devoted to students with depressive symptomatology. Although college students typically commit suicide at a lesser rate than do non-college students of a similar age, it is still a major concern for college mental health professionals. It is important to differentially diagnose depression from bipolar disorder as well as from a fairly large number of potential physical problems. During recent years, there has been an increased use of consultations with medical colleagues as this vital physical/psychological link has become more understood. Many mental health professionals now see that the prescription of psychotropic medication is an effective way in which to quickly handle the crisis experienced by severely depressed students. Usually, the medicating of depression is done in conjunction with psychological interventions, a strategy that appears to have the greatest ameliorative effect. The goal is to keep these students alive, to return them to an acceptable level of functioning, to help them learn more effective ways of handling stress (if it is more than simply biochemical), and to enable them to continue their educational pursuits.

2.3.5. Anxiety

An area of concern that often coexists with depression is anxiety. This may range from minor apprehensions about performance, appearance, and social acceptability

to panic attacks, phobias, and obsessive–compulsive disorders. Although not as life-threatening as depression, anxiety can still become life-disabling. Unfortunately, students with anxiety problems are often overlooked. Because everyone is anxious about something—especially in school—its very commonality means that it can be an invisible plague, damaging performance and interfering with students' quality of life. Evaluation of anxious students involves an exploration of current and past stressors, a personal and family history of anxiety or other psychological issues such as depression, and the current or past use of medication. Anxious students may present themselves with a variety of symptoms such as inability to concentrate while studying or sitting in class, difficulty in sleeping, avoidance of social situations, and compulsive performance of repetitive behaviors. Again, as with depression, coordination of treatment with medical professionals can be the most effective intervention.

2.3.6. Attention Deficit/Hyperactivity Disorder

Attention deficit disorders are relatively new areas of concern for college counselors. Although these issues have been around for as long as there have been learners, only recently have there been concerted efforts to respond appropriately to these problems. These are clearly dysfunctions in the neurophysiology and biochemistry of the individuals concerned. Unfortunately, these problems have been very difficult to measure and document, leaving sufferers with significant deficits and very little remediation. Adult attention deficit/hyperactivity disorder (ADHD) has only recently been clearly recognized, diagnosed, and medicated. Some of the signs include making careless mistakes, being unable to concentrate, not seeming to listen, often losing things, being easily distracted by external stimuli, and being forgetful. Many students who are evaluated as having ADHD have at least one parent who had the same problems and so they thought it was normal. Others have had ADHD identified all the way through school but have parents who have been reluctant to use medication due to side effects, stigma, or disbelief that their children had anything wrong with them. For many of those who have ADHD, the use of medication can literally change their lives, enabling them to concentrate and to actually enjoy learning for the first time.

2.4. Environmental Problems

External events may affect and interfere with students' academic progress. The accidental or intentional deaths of fellow students can cause distress among a large number of students, even if they are not acquainted with those who have died. Special counseling services dealing with bereavement issues may be necessary for a large number of students at one time. Fear resulting from a series of physical assaults and/or robberies on campuses is a realistic response, yet it is one that affects students' mental functioning.

Rejection of individuals due to their lifestyles (e.g., gay, lesbian, bisexual, transgender), religion, race, or native countries on the part of a large number of students, faculty members, and/or administrators often causes disruptions in students' efficiency. Students facing the decision to "come out"—and to whom—are often conflicted about the best course of action. Recently, the invasion of Iraq caused many Muslim students to fear for their safety on campus. Racial tension on campus often is the result of one group being perceived as receiving more support from the university. Individual or group counseling can be an important resource during these times.

3. COUNSELING RESOURCES

Recognizing the large number of adjustment demands on students, a variety of counseling services are supported by postsecondary institutions. Sources of counseling include individuals, agencies, and organizations. Colleges often employ academic advisers, psychologists, and psychiatrists. In addition, campus ministries have people trained in pastoral counseling, and they are a wonderful source for personal and religious counseling. Resident hall assistants who, like the residents, are also students are a source of peer counseling. Many counseling centers hire counseling psychologists to work with college-age students in the development of more appropriate adjustment strategies and in the alleviation of stress-related symptoms. Agencies such as career service centers, ombudsperson's offices, and student health centers offer specialized counseling that may be preventive and/or remedial in nature. Postsecondary institutions recognize the need for, and their responsibilities in offering, counseling services to their students.

See Also the Following Articles

Alcohol Dependence ■ Attention Deficit/Hyperactivity Disorders (ADHD) ■ Career Counseling ■ Diagnostic and Statistical Manual of Mental Disorders ■ School-to-Work Transition ■ Stress ■ Vocational Psychology, Overview

Further Reading

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Benton, S. A., Robertson, J. M., Tseng, W., Newton, F. B., & Benton, S. L. (2003). Changes in counseling center client problems across 13 years. *Professional Psychology: Research and Practice*, *34*, 66–72.

Erikson, E. H. (1968). *Identity: Youth and crisis*. New York: Norton.

Gallagher, R. P. (2003). *National Survey of Counseling Center Directors*. Alexandria, VA: International Association of Counseling Services.

Garfield, S., & Bergin, A. (Eds.) (1986). *Handbook of psychotherapy and behavior change* (3rd ed.). New York: John Wiley.

Holmes, T. H., & Rahe, R. H. (1967). The Social Readjustment Rating Scale. *Journal of Psychosomatic Research*, *11*, 213–218.

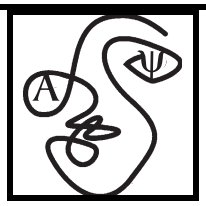
Marcia, J. E. (1980). Identity in adolescence. In J. Adelson (Ed.), *Handbook of adolescent psychology* (pp. 159–187). New York: John Wiley.

Maslow, A. H. (1970). *Toward a psychology of being* (2nd ed.). New York: Harper & Row.

Mental health: Does therapy help? (1995, November). *Consumer Reports,* pp. 734–739.

Rathus, S. A., & Nevid, J. S. (1998). *Adjustment and growth: The challenges of life* (7th ed.). Fort Worth, TX: Harcourt Brace.

Weiten, W., & Lloyd, M. A. (2000). *Psychology applied to modern life: Adjustment at the turn of the century* (6th ed.). Belmont, CA: Wadsworth/Thomson Learning.



Posttraumatic Disorders

David S. Riggs and Edna B. Foa

University of Pennsylvania, Philadelphia, Pennsylvania, USA

1. Definition of Trauma
2. Common Reactions to Trauma
3. Course of Recovery from Trauma
4. Psychological Disorders Following Trauma
5. Assessment in the Wake of Trauma
6. Psychological Treatment of Posttraumatic Disorders
Further Reading

GLOSSARY

acute stress disorder (ASD) A psychiatric disorder that develops in some people following a trauma; symptoms include dissociation, reexperiencing of the trauma, avoidance of trauma-related cues, and increased arousal.

arousal The tendency for individuals who have experienced a trauma to have elevated levels of physiological and emotional tension; this may be manifest as increased agitation, irritability, and/or startle as well as difficulty in concentrating or sleeping and in a sense of hypervigilance.

avoidance The tendency for individuals who have experienced a trauma to stay away from reminders of the trauma.

numbing A sense of diminished emotional responding commonly seen in trauma survivors; it is sometimes described as an empty or blank feeling.

posttraumatic stress disorder (PTSD) A psychiatric disorder that develops in some people following a trauma; symptoms must persist for at least 1 month and include reexperiencing of the trauma, avoidance of trauma-related cues, and increased arousal.

psychological trauma The emotional impact of a traumatic event on the person who experienced it.

reexperiencing The tendency for memories, thoughts, feelings, and/or images related to a past trauma to intrude into the current day; manifestations of reexperiencing symptoms include intrusive thoughts or images, nightmares, and extreme emotional and physiological distress when reminded of the trauma.

traumatic event An event that involves actual or threatened death or injury or a threat to the physical integrity of oneself or others.

traumatized people People who have experienced a traumatic event that resulted in psychological trauma.

The psychological aftereffects of traumatic events can be dramatic, and in some cases they can be prolonged. There are a group of common psychological reactions to traumatic events that reflect the overwhelming nature of the trauma. These reactions, when severe or prolonged, may develop into the psychological disorders of posttraumatic stress disorder (PTSD) and acute stress disorder (ASD). This article describes common reactions to trauma, trauma-related psychiatric disorders, common issues in the assessment of trauma and posttrauma symptoms, and treatments for posttraumatic stress disorders.

Psychological reactions following traumatic events have been recognized for hundreds (or perhaps thousands) of years in literary descriptions and for more than 100 years in the psychiatric field. Terms such as war neurosis, shell shock, and battle fatigue reflect the fact that these reactions were largely associated with war. More recently, researchers and

©2004 Elsevier Inc.
All rights reserved.

clinicians have begun to examine the psychological impact of a wide range of traumatic events, including rape, physical assault, natural and manmade disasters, motor vehicle accidents, and chronic abuse in various populations such as child abuse, domestic violence, and political imprisonment. This research clearly demonstrates that traumatic events can have devastating and long-lasting effects on the psychological well-being of survivors. At the same time, this research also indicates that humans are remarkably resilient and that, with time, most trauma survivors recover on their own.

1. DEFINITION OF TRAUMA

Individuals will label many different events as traumatic. The formal definition of a traumatic event adopted by the American Psychiatric Association in the fourth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV) consists of two distinct aspects. The first describes the characteristics of the event itself, and the second describes the subjective emotional responses of the individual who experiences the event:

1. The person has experienced, witnessed, or been confronted by an event that involved actual or threatened death or injury or a threat to the physical integrity of oneself or others.
2. The person responds to this event with intense feelings of fear, horror, or helplessness.

The definition allows different types of events to be identified as traumatic and recognizes that some people will experience psychological trauma from a particular event, whereas others will not. Although experiencing traumatic events was once thought to be rare, it is now known that most people are exposed to at least one traumatic event over the course of their lifetimes.

2. COMMON REACTIONS TO TRAUMA

The psychological impact of a traumatic event differs across types of events and individuals. Some events (e.g., rape, combat) are more likely to result in long-lasting psychological problems than are others (e.g., motor vehicle accidents, natural disasters). Also, individuals who experience the same event may differ greatly in their reactions to it. Despite this variation, there is a cluster of reactions that commonly occur following a trauma. Table I presents a group of reactions that many trauma survivors experience. Research on recently traumatized people suggests that nearly all trauma survivors will experience at least some of these reactions. These include symptoms that reflect elevated levels of emotional distress, such as increased feelings of fear and anger and general increased arousal, as well as feelings of guilt, shame, and grief. In addition, shortly after a traumatic event, people often find themselves replaying or reliving the event in their minds. These intrusive thoughts are usually accompanied by feelings similar to those experienced at the time of the traumatic event. The distress associated with the traumatic memories often leads trauma survivors to avoid people or situations that remind them of the traumatic event.

3. COURSE OF RECOVERY FROM TRAUMA

Shortly after a traumatic event, most traumatized people experience significant emotional distress that, in the majority of cases, diminishes over time. Although there are individual differences in patterns of emotional recovery from a traumatic event, most people recover within the first few weeks or months after the event. However, a substantial minority experience persistent chronic psychological problems related to the trauma. In some cases, individuals appear to recover relatively well from the trauma, only to have trauma-related problems surface months or years later.

Several factors have been found to distinguish people at most risk for persistent trauma-related problems, particularly those people who develop chronic posttraumatic stress disorder (PTSD). The factors that appear to be related to persistent posttraumatic problems can be loosely divided into (a) characteristics of the traumatic event, (b) characteristics of the individual exposed to the trauma, and (c) characteristics of the environment in which the individual is attempting to recover.

Among the characteristics of the traumatic event that are associated with persistent posttrauma problems is the type of trauma, with sexual assaults and combat being more likely to result in prolonged problems than nonsexual assaults and motor vehicle accidents. Exposure to a more severe traumatic event

TABLE I
Common Psychological Reactions to Traumatic Events

Reaction	*Examples*
Fear and anxiety	Fear of a repetition of the trauma
	Fear in response to environmental cues reminiscent of the trauma such as people, places, times of day, smells, and noises
	Anxiety attacks that feel as though they come out of the blue
	General sense of dread or anxious arousal
Reexperiencing of the trauma	Unwanted thoughts of the trauma
	Unwanted mental images of the trauma
	Flashbacks in which the person feels the trauma is occurring again
	Bad dreams or nightmares
Increased arousal	Feels jumpy, jittery, or shaky
	Feels irritable or angry
	Has difficulty in concentrating
	Has difficulty in sleeping
Avoidance	Attempts to suppress trauma-related thoughts and feelings
	Avoids trauma-related conversations
	Avoids situations or activities that are reminiscent of the trauma
Numbing	Sense of emotional numbness or emptiness
	Detachment from surroundings and other people
	Depersonalization or derealization
Guilt and shame	Blames self for things he or she did or did not do
	Is ashamed of what happened during the trauma or by the nature of the traumatic event
	Is ashamed or guilty about the emotional or behavioral reactions the person experienced during and after the trauma
Grief and depression	Feels down, sad, hopeless, or despairing
	Has loss of interest in, and decreased enjoyment from, normal activities
Altered image of self and the world	Sees the world as generally or uniformly dangerous
	Views the self as incompetent or bad in some way
	Views others more negatively and less trustworthy
	Is concerned that the person's reactions to the trauma mean that he or she is losing control or "losing my mind"
Disrupted relationships	Has difficulties with intimate relationships
	Feels distanced or cut off from friends and family
Alcohol and substance use	Increased use of alcohol and drugs to "self-medicate" reactions to the trauma

is also more likely to produce persistent problems. Markers for a severe trauma include physical injuries and the duration of the event. Individual characteristics that are associated with persistent problems include prior exposure to traumatic events and the presence of diagnosable depression or anxiety prior to the event. Among posttrauma factors, the absence of a supportive social environment appears to be associated with persistent problems following a trauma.

Additional markers that predict persistent emotional distress reflect the interaction of trauma and individual characteristics. For example, the perception that one is likely to die during the trauma, which reflects both the severity of the trauma and the individual's interpretation of the events of the trauma, is associated with persistent posttrauma problems. Similarly, intense emotional reactions (e.g., fear, helplessness, shame, anger) during and shortly after the trauma, as well as severe initial posttrauma reactions (e.g., intrusive memories, increased arousal, dissociation), are associated with more persistent problems.

4. PSYCHOLOGICAL DISORDERS FOLLOWING TRAUMA

4.1. Posttraumatic Stress Disorder

PTSD was first introduced into the diagnostic nomenclature in the third edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-III) in 1980. PTSD is classified as an anxiety disorder, but it incorporates three distinct groups or clusters of symptoms: reexperiencing, avoidance, and increased arousal. The reexperiencing cluster consists of five symptoms that reflect distressing spontaneous or cued recollections of the traumatic event such as intrusive memories, nightmares, and emotional distress in response to trauma-related cues. The seven avoidance symptoms reflect attempts at avoiding memories or reminders of the event as well as disruptions in emotional/interpersonal experience such as a numbing of emotional reactivity and a sense of detachment from other people. The increased arousal cluster consists of five symptoms that reflect an elevated level of alertness (e.g., hypervigilance, increased startle) and a high level of general arousal (e.g., sleep disturbance, irritability). To be diagnosed with PTSD, an individual exposed to a traumatic event (as defined by the DSM-IV) must experience symptoms in each of these clusters. Specifically, the person must manifest at least one reexperiencing, three avoidance, and two increased arousal symptoms for at least 1 month after the traumatic event. If the symptoms persist for more that 3 months, the condition is considered chronic.

4.2. Acute Stress Disorder

Acute stress disorder (ASD) was introduced as a distinct disorder in the DSM-IV. ASD is characterized as an anxiety disorder and shares many symptoms in common with PTSD. One important difference between the two disorders is the time frame of the symptoms. The diagnosis of ASD requires that the symptoms occur within the first 4 weeks after a traumatic event, whereas PTSD requires that the symptoms persist for at least 1 month after the traumatic event. The other feature that distinguishes the two disorders is a greater emphasis on dissociative symptoms in ASD than in PTSD.

To be diagnosed with ASD, a person must be exposed to a traumatic event and develop symptoms in each of four categories: dissociation, reexperiencing, avoidance, and arousal. In contrast to PTSD, the avoidance cluster of ASD symptoms consists of only effortful avoidance of trauma reminders. Symptoms of emotional numbing and amnesia are included in the dissociation cluster, and other symptoms (e.g., anhedonia, foreshortened future) are not included. In contrast to the diagnosis of PTSD, ASD requires only one symptom in each of the reexperiencing, avoidance, and arousal clusters.

Initially, it was thought that that the presence of ASD would identify people who would go on to develop chronic PTSD. However, recent research suggests that the relationship between acute reactions to trauma and more chronic problems is not so straightforward. ASD does appear to serve as a risk marker for PTSD with many people who develop chronic PTSD meeting criteria for ASD shortly after a trauma. However, some survivors with ASD do not go on to develop PTSD, and some traumatized people who do not have ASD may develop PTSD later. The imperfect relationship between the diagnoses of ASD and PTSD probably reflects the fact that the initial severity (not simply the presence or absence) of symptoms following a traumatic event is related to their persistence. Also, relating the presence of ASD directly to a PTSD diagnosis ignores factors that augment or impair recovery during the first month following a traumatic event. Regardless of its utility as a precursor to PTSD, the ASD diagnosis does capture the significant distress that some people experience shortly after trauma—distress that cannot be diagnosed as PTSD because the symptoms have not yet persisted for more than 1 month (Table II).

5. ASSESSMENT IN THE WAKE OF TRAUMA

Assessments of posttrauma pathology focus on three areas: (a) exposure to traumatic events, (b) trauma-related symptoms, and (c) establishment of the link between the traumatic event and the symptoms. Establishing the link between the event and the symptoms is particularly difficult in cases of chronic PTSD where the traumatic event may have occurred years or even decades in the past and in cases where the patient has experienced multiple traumas. Establishing this link is typically less complicated when assessing ASD where the traumatic event must have occurred recently.

TABLE II
Symptoms of Posttraumatic Stress Disorder and Acute Stress Disorder

Posttraumatic Stress Disorder	*Acute Stress Disorder*
Reexperiencing symptoms (need at least one)	*Dissociation (need at least three)*
1. Intrusive thoughts or memories of the trauma	1. Subjective sense of emotion numbing or detachment
2. Recurrent dreams related to the trauma	2. Reduced awareness of surroundings
3. Flashbacks of the trauma	3. Derealization
4. Emotional distress to trauma reminders	4. Depersonalization
5. Physiological reactivity to trauma reminders	5. Unable to recall important aspect of the trauma
Avoidance (need at least three)	*Reexperiencing (need at least one)*
6. Efforts to avoid thoughts or feelings of trauma	6. Recurrent trauma-related images, thoughts, dreams, or flashbacks
7. Efforts to avoid trauma reminders	7. Emotional distress to trauma reminders
8. Unable to recall important aspect of the trauma	*Avoidance (need at least one)*
9. Diminished interest in activities	8. Marked avoidance of trauma-related thoughts, feelings, or reminders
10. Detachment from others	*Arousal (need at least one)*
11. Restricted range of affect	9. Marked symptoms of arousal such as sleep problems, irritability, difficulty in concentrating, hypervigilance, exaggerated startle, restlessness
12. Foreshortened sense of the future	
Arousal (need at least two)	
13. Sleep disturbance	
14. Irritability or angry outbursts	
15. Difficulty in concentrating	
16. Hypervigilance	
17. Increased startle response	

5.1. Assessing Exposure to Trauma

The specific guidelines for assessing exposure to traumatic events depend on the goals of the assessment. This being said, five aspects of traumatic exposure are typically assessed during the evaluation: (a) establishing the presence of trauma in the life of the client, (b) identifying the nature of the traumatic exposure, (c) estimating the severity of the trauma, and (d) establishing the timing of the traumatic exposure.

5.1.1. Establishing the Presence of Trauma

Typically, the assessment of traumatic exposure includes an evaluation of traumatic events across the life span. In some cases, the primary focus may be on a single event, but understanding the prior trauma history of the person who experienced the trauma is important.

The presence of traumatic exposure is typically determined by presenting the person with a list of traumatic events and having the person indicate those events to which he or she has been exposed. The specific events included on the list will vary across settings, situations, and clinicians. The most commonly evaluated traumas include motor vehicle accidents, physical and sexual assaults, physical and sexual abuse (childhood or adult), fires, natural disasters, and combat. It is better to use structured questions that describe each event in detail rather than a single question such as "Have you experienced a trauma?" This helps to ensure that the person's response does not rely on his or her understanding of a single word such as "trauma." Similarly, the wording of each item should be designed to reduce the likelihood of misunderstanding or misinterpretation. For example, it is better to ask "Has anyone ever used physical force or threats of force to make you do something sexual that you didn't want to do?" instead of "Have you ever been raped?" There are many times when a person reports having experienced forced sexual contact but will not view it as rape.

In addition to determining what types of traumatic experiences a person has experienced, assessments should inquire about the number of times various types of events have occurred. For example, the assessor will want to know how many earthquakes the person has experienced or how many times the person has been assaulted.

5.1.2. Identifying the Nature of the Traumatic Exposure

Although traumatic events are thought to share certain characteristics, it is clear that they are not the same in their emotional impact. A more complete understanding of the traumatic event and its potential impact on a person requires a thorough understanding of the specific characteristics of the event and the nature of the person's exposure to that event. The definition of trauma included in the diagnostic criteria for PTSD and ASD allows for three types of exposure to the event: (a) directly experiencing the event, (b) witnessing the event, and (c) being confronted with the event. In general, but not always, directly experiencing a trauma is more likely to produce lasting effects on the person than is witnessing the event or being confronted by it.

Often, a distinction is made between interpersonal traumas (e.g., criminal assaults, child abuse) and other types of traumas (e.g., fires, natural disasters). Research suggests that interpersonal traumas are more likely to produce persistent reactions than are other types of traumatic events. However, this might not be true in all cases.

Traumas have also been differentiated by whether the event happened to a particular individual or to a group of people. Events such as criminal assault, child abuse, and life-threatening illnesses typically happen to only one person. In contrast, events such as natural disasters, wars, and the terrorist attacks of September 11, 2001, affect entire communities simultaneously. Although individual reactions may be quite varied, there is some suggestion that individual traumas are more likely to result in chronic problems than are community traumas. However, the distinction between individual traumas and community traumas is far from perfect. For example, a violent crime may typically be perpetrated against an individual, but the cumulative effect of crime can affect an entire community. Similarly, although a hurricane or another natural disaster may affect an entire community, the specific experiences of individuals within that community may be quite varied.

5.1.3. Assessing the Severity of the Traumatic Event

Two aspects of trauma severity are typically assessed: objective severity and subjective severity. These map loosely onto the two aspects of the DSM-IV definition of trauma in the diagnostic criteria for PTSD and ASD. Measures of objective dangerousness will depend on the specific type of traumatic event, but they reflect characteristics of the event such as whether someone was killed, whether and how badly the person was injured, and whether an assailant used a weapon during the assault. Other objective measures of trauma severity might include the duration of the event and the amount of monetary or property loss.

Subjective severity of the trauma refers to the person's evaluation of the event and overlaps considerably with the second part of the trauma definition mentioned previously. Therefore, these estimates of severity tend to focus largely on feelings of fear, horror, and helplessness. Typically, survivors are asked to what degree they felt each of these emotions during the traumatic event. For diagnostic purposes, all that the assessment needs to establish is that the person experienced the emotion, and simple questions such as "Did you fear that you would be killed during the trauma?" are sufficient. In other cases, one may wish to establish the subjective severity relative to other events. In these cases, it is often helpful to have the person rate the intensity of the feeling on an arbitrary scale.

The objective and subjective ratings of trauma severity may be quite different. Seemingly minor (based on objective ratings) events are rated as subjectively very severe and vice versa. Therefore, it is important and informative to assess both the objective severity and subjective severity of traumatic events.

5.1.4. Assessing the Timing of the Traumatic Event

Timing of the trauma refers to both the developmental level (i.e., age at which the event occurred) and the duration over which the event occurred. In this case, duration usually refers to whether the event occurred a single time (e.g., assault by a stranger on the street) or was repeated over time (e.g., physical abuse by a parent over years). There is some indication that exposure to a traumatic event at an early age or repeated exposure to a traumatic event over a period of time is more likely to lead to psychological problems than is a single traumatic event during adulthood. Some have argued that these problems extend beyond those captured by the PTSD diagnosis and include disrupted sense of self and relationships.

5.2. Assessing Early Emotional Reactions to the Traumatic Event

As described earlier, it is common for people who have experienced a traumatic event to report reactions such

as intrusive memories, emotional distress, and increased arousal. In most situations, there will be little formal assessment of these reactions unless the person is being seen by a mental health professional. Informal assessment of these reactions might consist of a review of the common reactions listed in Table I. Such informal assessments respect the survivor's wish to limit what he or she discusses, and the presence of the common reactions discussed previously should be normalized and validated. Because these reactions will diminish over time naturally, it is recommended that a recently traumatized person be encouraged to use his or her naturally occurring support system and to follow his or her natural inclination with regard to how much to say and to whom to talk. However, an individual who experiences very severe emotional reactions or who shows high levels of dissociative symptoms should be referred for further evaluation and possible treatment.

5.3. Assessment of Symptoms

Two areas are typically evaluated in a comprehensive assessment of trauma-related symptoms. The first is the specific symptoms included in the PTSD or ASD diagnostic criteria. The second involves the assessment of associated symptoms and problems. As with the assessment of trauma exposure, the specific questions or instruments included in the assessment battery will depend largely on the purpose of the assessment and the specific trauma history of the person being evaluated.

5.3.1. Assessing the Symptoms of PTSD or ASD

Both self-report and interview-based instruments have been developed to assess the symptoms of PTSD and ASD. In some cases, only the presence or absence of each symptom is assessed, but most instruments also call for rating the severity of the symptoms. The specific metric on which the severity of the symptoms is measured varies across instruments, but usually the measurement incorporates an evaluation of the frequency with which each symptom has occurred and the intensity or severity of each symptom when it is experienced. Typically, PTSD symptoms are assessed over a 1-month time frame, reflecting the criteria that symptoms must be experienced for at least 1 month to be diagnosed with the disorder, but symptoms have been assessed over periods as short as 1 week when necessary. Because ASD can occur only within the month immediately following a trauma, these symptoms are generally evaluated relative to a time frame of 1 week or even a few days.

When assessing symptoms of PTSD or ASD, it is important to relate the symptoms to a specific traumatic event and to determine whether the current presentation reflects a shift from prior functioning. In the case of some of the symptoms, this is relatively straightforward because the symptoms are directly tied to the event (e.g., intrusive thoughts or memories of the event, efforts to avoid reminders of the event). In the case of other symptoms, however, the association with the specific traumatic event in question is not necessarily clear (e.g., sleep disturbance, sense of detachment from other people). In the latter case, the evaluator will generally find it necessary to use the "change from prior functioning" criteria to link the symptoms to a particular trauma. As noted earlier, this task is easiest when the traumatic event in question is relatively recent and discrete. When the trauma occurred a long time ago (particularly when it occurred during childhood) or over a long period of time (as in the case of domestic violence), it is much more difficult to establish the link to a particular event or a shift from prior functioning.

5.3.2. Assessing Broader Difficulties Associated with Trauma

Beyond the symptoms of PTSD and ASD, trauma survivors can experience a wide range of problems, including (a) mental health problems such as depression, anxiety, and substance abuse/dependence; (b) problems in functioning such as disrupted relationships, difficulty in working, and physical health problems; and (c) cognitive and emotional difficulties such as intense feelings of guilt or shame, anger control problems, and problem-solving deficits. Some instruments have been developed to assess these broader difficulties associated with trauma, but for the most part, evaluators have used established instruments for assessing each of these areas of interest.

6. PSYCHOLOGICAL TREATMENT OF POSTTRAUMATIC DISORDERS

Psychological treatments for trauma survivors can be divided into those designed to be implemented in the immediate aftermath of the traumatic event (often termed "debriefing"), those aimed at treating ASD,

and those aimed at treating chronic PTSD. Treatments aimed at alleviating PTSD have been studied much more extensively than have those used for ASD or debriefing.

6.1. Treatment of Chronic PTSD

The most well-studied and empirically validated treatments for PTSD are based on behavioral and cognitive–behavioral approaches to therapy. Three major categories of treatments have been tested and found to be somewhat effective for patients with PTSD: (a) exposure therapy in which patients are encouraged to systematically confront reminders of the traumatic event that elicit anxiety, (b) cognitive techniques in which patients are taught to challenge cognitive distortions thought to underlie the symptoms of PTSD, and (c) anxiety management techniques in which patients are taught methods for managing their trauma-related anxiety. In addition, various combinations of these treatments have been examined and found to be useful in treating chronic PTSD.

Of the treatments used for chronic PTSD, exposure therapy has been the most widely validated. Two forms of exposure are typically included in these programs: imaginal and in vivo. Imaginal exposure refers to patients intentionally confronting specific memories of the traumatic event in their imaginations either by verbally describing the traumatic event or by writing a description of the trauma. Typically, this is done for a prolonged period of time (i.e., 30–60 minutes) and is repeated over the course of treatment. The second type of exposure, in vivo exposure, refers to real-life confrontation with situations that are objectively safe but elicit intense anxiety or with situations that elicit memories of the trauma itself. The most effective therapies for chronic PTSD include these two exposure techniques or combine one or both of these techniques with cognitive therapy techniques aimed at altering PTSD-related dysfunctional cognitions.

6.2. Treatments for ASD

Treatments for ASD have been studied much less extensively than have treatments for chronic PTSD. The available studies suggest that the cognitive–behavioral techniques found to be useful with cases of chronic PTSD are also helpful in the more immediate aftermath of trauma. These interventions are typically brief (four or five sessions, with each session lasting 90–120 minutes) and are conducted during the first 2 to 4 weeks after the trauma. Despite the early promise of these interventions, it is premature to conclude that treatment targeting ASD symptoms during the first month after a traumatic event are effective in reducing long-term difficulties associated with trauma.

6.3. Debriefing: Treatment in the Wake of Trauma

Although debriefing interventions (typically administered in a single session during the first several days after a trauma) have become widely used, they have been poorly studied. These interventions aim to reduce long-term mental health problems associated with the trauma, but there is little evidence to support their efficacy in achieving this goal. Indeed, there is growing evidence that debriefing does not substantially reduce the rates of PTSD among trauma survivors. Of more concern are recent findings suggesting that debriefing might actually be detrimental to some trauma survivors, leading to more long-term emotional disturbance than would have occurred with no intervention at all.

See Also the Following Articles

Stress

Further Reading

American Psychiatric Association. (1994). *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.). Washington, DC: American Psychiatric Press.

Bryant, R. A., & Harvey, A. G. (2000). *Acute stress disorder: A handbook of theory, assessment, and treatment*. Washington, DC: American Psychological Association.

Davidson, J. R. T., & Foa, E. B. (1993). *Posttraumatic stress disorder: DSM-IV and beyond*. Washington, DC: American Psychiatric Press.

Foa, E. B., Keane, T. M., & Friedman, M. J. (2000). *Effective treatments for PTSD: Practice guidelines from the International Society for Traumatic Stress Studies*. New York: Guilford.

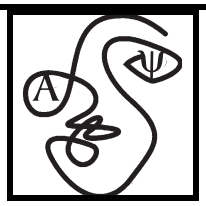
Follette, V. M., Ruzek, J. I., & Abueg, F. R. (1998). *Cognitive–behavioral therapies for trauma*. New York: Guilford.

Saigh, P. A., & Bremner, J. D. (1999). *Posttraumatic stress disorder: A comprehensive text*. Boston: Allyn & Bacon.

Wilson, J. P., Friedman, M. J., & Lindy, J. D. (2001). *Treating psychological trauma and PTSD*. New York: Guilford.

Wilson, J. P., & Keane, T. M. (1997). *Assessing psychological trauma and PTSD*. New York: Guilford.

Yehuda, R. (1998). *Psychological trauma*. Washington, DC: American Psychiatric Press.



Power, Authority, and Leadership

Lourdes Munduate and Francisco J. Medina
University of Seville, Seville, Spain

1. Power and Influence
2. Formal Authority and Leadership
3. Models of Leadership
Further Reading

GLOSSARY

authority The institutionalized power between a superior and a subordinate that ensures compliance with the superior's wishes because he or she is the boss.

empowerment A motivational process involving feelings of self-efficacy among organizational members; to enhance the perception of self-efficacy, members use a set of strategies to influence and improve their behavior, taking responsibility for their own lives rather than depending on a leader to direct and motivate them.

influence A force that one person (the leader) exerts on others (followers) to induce a change in the latter, including changes in behaviors, attitudes, and values.

leadership Power relationships and influence processes between a leader and followers in which the leader exercises greater influence over the followers to accomplish group or organizational objectives.

power The potential ability of a person to influence others; thus, influence is "power in action," just as power is "potential influence."

Leadership involves power relationships and influence processes between a leader and followers in which the leader exercises greater influence over the followers to accomplish group or organizational objectives. Thus, power is essential to leadership and processes of influencing members of a group or an organization. However, although power is necessary for leadership, it is insufficient by itself for leadership. The same person could be described as having influence, being persuasive, and/or being a leader. As a result, social scientists have attempted to discover how leaders use power to influence followers, looking at the difference between exercising formal authority and leadership and followers' responses to these influence attempts. It should be noted that there is a clear difference between yielding to social pressure from an individual and being genuinely persuaded. This article first examines the various bases and sources of power that are used to influence people and the responses they elicit in followers as well as the difference between management and leadership. Then, it focuses on three fundamental features that are critical to leadership: traits or personal characteristics, behaviors, and situations. Finally, it describes some integrative approaches represented mainly by charismatic and transformational leadership.

1. POWER AND INFLUENCE

In trying to understand how people influence each other when they interact, social psychologists have used the concepts of social power and influence tactics. The two terms have sometimes been interchanged, but a clear distinction needs to be made between them.

©2004 Elsevier Inc.
All rights reserved.

Social power reflects the repertoire of tools available to a person to influence the environment or the other party. Influence tactics refer to the actual use of a certain tool in a specific situation. The definition of the two concepts helps to clarify the differences between them. Following Lewin's initial conceptualization in 1941, French and Raven in 1959 defined influence as a force that one person (the leader) exerts on others (followers) to induce a change in the latter, including changes in behaviors, attitudes, and values. Social power was subsequently defined as the potential ability of a person to influence others. Thus, influence is "power in action," just as power is "potential influence."

The conceptualization of power highlights two important characteristics of power. The first reflects that power is potential and, as a result, that a person can have power without using it. The availability of tools does not mean that the tools must be used. The second characteristic refers to the relational aspect of power in that it does not reside in the person but rather resides in the relationship between people in a given situation. In general, power is derived from the situation in which the person operates, and in the event of the situation changing, the person's power also changes.

1.1. Social Influence Dynamic

In studies examining people's responses to leaders' influence attempts, there is a clear difference between yielding to direct or indirect social pressure from a leader and being genuinely persuaded. For example, the leader's influence may be strong enough to wield control over the followers' behavior, ensuring public agreement, regardless of whether the followers are privately convinced. On other occasions, the influence process may change followers' private attitudes or opinions and make them committed to the leader's request. Therefore, social psychologists have proposed various types of influence processes, the most important of which was Kelman's traditional distinction, made in 1958, among compliance, internalization, and identification. Compliance refers to a surface change in behavior and expressed attitudes, often as a consequence of coercion or the followers' desire to obtain a reward. Because compliance does not reflect internal change, it usually persists only while behavior is under surveillance. In contrast to compliance, internalization means subjective acceptance that produces true internal change that persists in the absence of surveillance. The norm becomes an internalized standard for behavior. There is genuine support for the leader's proposals because they appear to be intrinsically desirable and are suited to the followers' values and beliefs. Identification is based on the leader's attractiveness to the followers, who imitate the leader's behavior and attitudes to gain his or her approval. The maintenance of close relationships with the leader involves the followers' need for acceptance and esteem.

The outcomes of these various forms of social influence in terms of compliance, internalization, and identification may occur at the same time. For example, a person who works in a nongovernmental organization (NGO) with exploited children in the Third World may accept a proposal from a superior to design a new program because he or she will obtain tangible rewards for carrying it out (compliance) and believes in the necessity and effectiveness of the program (internalization). At the same time, the person is proud of working with his or her superior on this project, and this will increase the level of satisfaction in the relationship with the superior (identification).

1.2. Bases and Sources of Power

To understand the different ways in which power can be exercised, it is best to look at the various bases and sources of power. In 1959, French and Raven proposed a well-known taxonomy of power that considered five bases of power: reward, coercion, legitimacy, reference, and expertise (a sixth base of power, information, was added a little later). A leader possesses these resources and can use them to change the beliefs, attitudes, and/or behaviors of followers. Reward and coercive power rely on others believing that the leader can provide them with the desired rewards or can punish them, respectively. Using either of these bases will induce only a superficial change in the followers. In other words, none of the followers' privately held beliefs, attitudes, or values is changed. Only compliance is obtained, and the continuation of this compliance depends on successful surveillance of the followers by the leader. A person possesses legitimate power when others believe that he or she has a legitimate right to exert influence over them and they are obliged to accept this influence. It cuts across all three types of influence processes and, as a result, may share elements of each of them. Referent power refers to followers' identification with the leader. It leads to private acceptance by followers by enabling them to maintain a satisfactory relationship with the leader and to see themselves as similar to the leader in certain

important areas. Expert power depends on followers' perception of having expertise or knowledge in a specific domain. If the followers perceive the leader as an expert, this will result in private acceptance on the part of the followers. Finally, informational power leads to internalized and lasting changes in the followers' beliefs, attitudes, and/or values. Compared with other social power bases, the changed behavior resulting from information is maintained without continued social dependence on the leader and is instead based on the perceived importance and validity of the information. Only informational power leads to cognitive change in followers because it immediately becomes independent of the leader.

Another related dimension of power refers to sources of power. When dealing with the bases of power, the focus was the repertoire of tools available to a person to influence the other party or the environment. In terms of sources of power, the focus points toward the way in which a person comes to control these tools. Two different sources of power have been identified: position and personal power. Position power arises from the formal position held in a group or an organization, whereas personal power arises from personal attributes and the kind of relationship established with the other party. In position power, potential control is derived from legitimate authority. In personal power, it stems from task expertise or the opportunity to access certain information that is important for other members of the group or organization.

So far, this article has identified two sources of power, six bases of power, and three influence processes. Figure 1 illustrates the relationships among these aspects of power.

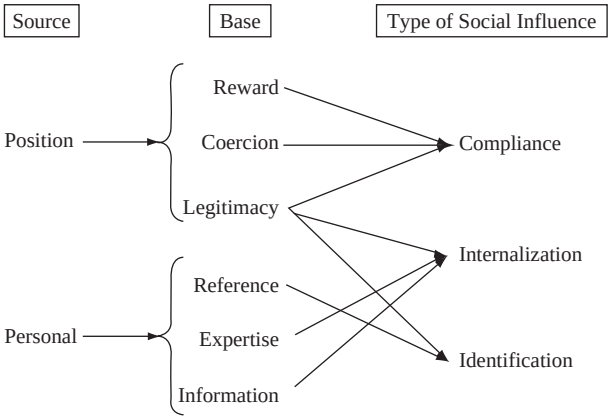


FIGURE 1 Relationships of sources, bases, and types of social influence elicited in followers.

1.3. Influence Tactics

As was explained earlier, literature on social power takes power to be the capacity to exert influence. However, power can be used in different ways, and the specific ways in which power is enacted involve influence behavior. With the publication of *The Powerholders* in 1976, Kipnis approached influence behaviors from the perspective of the agent of influence. He stated that the choice of a means of influence depends mainly on the resources at the agent's disposal (i.e., on the power bases that the target possesses), on the target's inhibition to invoke a power base (as determined by the target's estimate of the costs of using a power base, the target's subjective values and attitudes, individual differences, and social norms), and on the resistance that one expects from the target toward the influence attempt. Subsequent developments by Kipnis and colleagues in 1980 identified the specific behaviors that people have at their disposal for influencing others in organizations. The various categories of proactive influence behaviors that were proposed were referred to as "influence tactics."

Building on and extending the previous work, Yukl and colleagues developed a classification with nine proactive influence tactics: rational persuasion, inspirational appeals, consultation, ingratiation, personal appeals, exchange, coalition tactics, legitimating tactics, and pressure. In general, effective leaders use a combination of influence tactics that they select depending on their judgment of what is most appropriate for a given situation. As Yukl pointed out in his recent book titled *Leadership in Organizations,* a leader prefers to use tactics that are socially acceptable, feasible in terms of the leader's power in relation to the followers, likely to be effective for a particular objective, and unlikely to require much time, effort, or cost to the leader.

2. FORMAL AUTHORITY AND LEADERSHIP

As noted in the previous section, power that stems from a person's formal position in a hierarchical structure is sometimes called formal authority or legitimate power. Authority represents the institutionalized power between a superior and subordinates that ensures compliance with the superior's wishes because he or she is the boss. Thus, authority involves a

person's right to exercise control over other people or the environment. For example, a manager usually has the legitimate right to make a job-related request to subordinates, who in turn have the legitimate right to request job-related information and assistance from the manager. However, a manager may have a great deal of legitimate power due to his or her position but may have little, if any, expert power. As a result, followers may think that the manager does not understand the problem they must deal with but that they must do the work because they have been told to do so by the boss.

2.1. Management and Leadership

Although nobody has argued that managing and leading are equivalent, it has sometimes been assumed that managers' control over some important resources involves leadership processes. In fact, the exercise of power that stems from a person's hierarchical position may facilitate the emergence of leadership. But hierarchical position is simply a context from which leadership may or may not emerge. Leadership does not arise by virtue of occupying hierarchical roles. There are differences between formal role designations (which deal with formal aspects of the relationship) and the emergence of leadership (which deals with informal aspects of the relationship). Thus, the emergence of leadership involves more than subordination and is related to the kind, or depth, of social influence elicited from followers. People who achieve acceptable influence, such as internalization or identification, are the ones who are defined as leaders. Therefore, leaders can be identified because they demonstrate leadership. Countless presidents, kings, generals, and military figures who were chosen or appointed (bottom-up vs top-down), and who have power and authority, give few examples of behaviors that could be associated with actual leadership. If a person with formal authority (i.e., a manager) is able to induce internalization or identification, he or she would have managed to exercise leadership. If this person is not able to do so, he or she would have exercised authority. In turn, a person who develops a relation that induces internalization or identification despite not holding a hierarchical role (i.e., a nonmanager), is exercising leadership. In other words, a person can be a leader without managing, and a person can be a manager without leading. From this perspective, leading and managing are viewed as different processes, although leaders and managers are not considered to be different types of people.

TABLE I
Authority and Leadership as Results of Social Influence

Social influence	*Type of power*
Compliance	Formal authority (management)
Internalization	Leadership (transactional)
Identification	Leadership (transformational)

Nevertheless, it must be emphasized that some overlap does exist among the various types of social influence and that more than one influence process may occur at the same time, as stated previously. It may also be the case that the same type of influence attempts may produce different individual responses. As a result, not all theoretical perspectives share the same conceptualization on the exercise of power, authority, leadership, and management. In fact, the relationship among these concepts has been the subject of extensive debate. Table I presents a summary of the conceptualization developed in this article.

2.2. Empowerment

Leaders and managers often share power with followers or subordinates for a period of time to accomplish a variety of organizational tasks and objectives. This process is known as power delegation. In 1988, Conger and Kanungo proposed that empowerment differs from delegation in that the former is a motivational process involving feelings of self-efficacy among organizational members and is not temporally limited by the leader. To enhance the perception of self-efficacy, members use a set of strategies to influence and improve their behavior, taking responsibility for their own lives rather than depending on a leader to direct and motivate them. In conventional terms of social influence, empowerment can be understood as a process of identification focused toward the development of a shared vision among organizational members. The aim of the empowerment process, by strengthening individual and/or collective efficacy, is to enable organizational members to solve organizational problems arising from the turbulent environment and diversification of the workforce.

3. MODELS OF LEADERSHIP

As stated previously, a person cannot lead unless there are power relationships between the leader and the

followers. However, power by itself is not a sufficient condition for leadership. Various perspectives and models have approached the question of the variables that need to be considered to achieve the exercise of leadership. The conventional perspective emphasized the personal attributes and acquired behaviors that leaders bring to the situation as well as the situational forces that shape and determine effective leadership. Some new perspectives emphasize that leadership is not only a transactional relationship between the leader and the followers but also produces important effects in followers' value systems.

3.1. Traits, Behaviors, and Situational Models

Some of the earliest research to approach the subject of leadership consisted of trait or dispositional studies. Trait models focus on what a person brings to leadership activities. This approach emphasizes the personal attributes of a leader that are treated as innate and unique. Accordingly, attention has focused on certain combinations of personality traits (e.g., tolerance of stress, emotional maturity, self-efficacy), motives (e.g., power orientation, relatively strong motivation to achieve, need for affiliation), and skills (e.g., technical, conceptual, interpersonal) because these might be reflected in the achievement of leader status and as possible predictors of a leader's effectiveness in promoting group or organizational performance. However, the studies aimed at uncovering the characteristics constantly linked to leadership have been inconclusive, and no definitive predictors of effective leadership have been established.

Researchers became discouraged with the trait approach, and around 1950 they began to look at what leaders or managers do in their jobs. The behavioral models of leadership analyze what leaders or managers do with whatever personal traits, skills, or motivational capacities. The primary concern of behavioral models focused on developing methods to determine what leaders do and measure relevant dimensions of their behavior that might be linked to group effectiveness and subordinate satisfaction. Two main dimensions that reflected how subordinates perceived a manager's behaviors were identified. Consideration (or relationship-oriented) behaviors reflect the extent to which a leader shows concern for subordinates' satisfaction and well-being. Initiating structure (or task-oriented) behaviors reflect the degree to which a leader explains and defines the roles of a task, assigns subordinates to various task roles, controls subordinates' performance, and provides feedback to subordinates. Because these are two independent dimensions, the behavior of a leader may be described as a combination of both. According to this perspective, the ideal leader would be someone with high scores on both dimensions.

Both the trait and behavioral models focus on personal attributes and acquired behaviors that leaders bring to the situation and provide a framework for identifying and cultivating leaders in organizations. However, an important difference between the models lies in the fact that traits are considered innate in a person—a person either does or does not have the gift—whereas specific behaviors may be learned. A difference in the practical implications of the two models is that if the trait proposals were correct, leaders could be selected to hold formal posts in groups and organizations. However, if the behavioral models were correct, it would be possible to teach leaders, such that programs could be designed to train people to become effective leaders. The interesting aspect of this last perspective is that it could increase the number of available leaders.

During the 1960s, situational models supplanted universalistic approaches. Situational models focus on situational variables that moderate the relationship between the leader's attributes and behaviors and his or her effectiveness. Researchers became convinced that predicting the success of leadership was more than just a case of signaling certain preferable traits or behaviors. The impossibility of obtaining congruent empirical results gave rise to a new emphasis on situational factors. The relation between leadership style and efficacy indicated that Style X would be appropriate for Condition A, whereas Style Y would be appropriate for Condition B. However, it was one thing to assert that the efficacy of leadership depended on the situation, but managing to identify those situational conditions was something completely different. Attempts to isolate those conditions gave rise to a wide variety of situational models of leadership, all of which focused on explaining how aspects of the situation moderate the relationship between leaders' traits, behaviors, or both and the outcomes arising from a particular situation. A representative sample of these models is next described briefly.

The least preferred coworker (LPC) contingency model introduced by Fiedler in 1967 considered two types of leadership styles that were regarded as fixed or immutable dispositional traits: directive or task-oriented

style and nondirective or relationship-oriented style. The effectiveness of both styles depended on three situational variables that were identified as the leader–member relationship (e.g., the degree of confidence and commitment between the leader and the followers), task structure (e.g., the extent to which procedures for performing the task are established), and position power (e.g., control of resources such as money or information). Thus, when selecting leaders, this model suggests matching a person's leadership style with the situations. For example, if there is a group or situation in which the leader is trusted, the task is highly structured, and position power is good, it would be more appropriate to assign a task-oriented leader than to assign a relationship-oriented leader.

The popular life cycle model proposed by Hersey and Blanchard in 1977 identified four leadership styles—telling, selling, participating, and delegating—each of which is basically a mixture of task behavior and relationship behavior styles. The model viewed the "maturity" of followers—defined in terms of ability, motivation, and self-confidence—as the chief contingency variable to which one of these leadership styles had to be matched. Thus, followers with low maturity, as is often the case with new employees, would be most responsive to a telling leadership style (characterized as relatively high on task behaviors and low on relationship behaviors), whereas a delegating style (low in both task and relationship behaviors) would be more appropriate for followers with high maturity, as is usually the case with experienced employees. Therefore, the model states that the leader should adapt his or her leadership style to the maturity of the followers. The life cycle model was widely marketed and used in many leader training and development programs in organizations during the 1980s despite having received only partial support from empirical research.

According to the path–goal model developed by House in 1971, the leader's main role is to help his or her followers achieve goals and to provide sufficient leadership and support to guarantee that their goals are compatible with the overall objectives of the group or organization. To achieve these objectives, the leader must identify two critical situational variables (follower needs and task demands) and then choose an appropriate leadership style (consideration style or initiating structure style). The behavior of the leader is motivating for his or her followers in that, first, the satisfaction of the requirements of these followers depends on a good performance and that, second, these followers need to receive the leadership, support, and rewards that are vital for achieving a good performance.

The normative decision model built by Vroom and Yetton in 1973, and developed by Vroom and Jago in 1988, considers that decision making is the critical leadership activity with groups. The model went further than did previous proposals on decision procedures by specifying which procedure will be the most effective in each of several specific situations. Two intervening variables for measuring decision effectiveness were suggested: decision quality and decision implementation or acceptance by the parties involved in the decision. In turn, these variables are affected by the decision-making procedure used by the leader: two varieties of autocratic decision (AI and AII), two varieties of consultation (CI and CII), and one variety of joint decision making as a group (GII). Depending on seven key situational variables, such as the amount of relevant information possessed by the leader and followers and the likelihood that followers will accept an autocratic decision, the model provides a set of rules for identifying the best decision procedure in terms of decision quality, parties' acceptance, and costs of decision making. Vroom and Yetton developed a decision process flow chart, known as a decision tree, to help leaders identify the optimal decision procedure for each situation.

A more recent situational model is the cognitive resource model, introduced by Fiedler and Garcia in 1987, which focuses on the cognitive resources of the leader (intelligence and experience), leader behavior (directive leadership), and two aspects of the leadership situation (interpersonal stress and the nature of the group's task). Basically, the theory points out that leader intelligence contributes to group performance only when the leader is directive and the followers require guidance to perform the task effectively. However, perceived stress moderates the relation between intelligence and decision quality. Finally, perceived stress moderates the relationship between leader experience and performance. The general idea is that leader effectiveness is predicted by intelligence in low-stress conditions and by experience in high-stress conditions.

The final situational model of leadership considered here is the tridimensional leadership theory developed in 2003 by Yukl, who considered that leadership behavior can be described in terms of three broad categories: task oriented (primary effect on efficiency and reliability), relations oriented (primary effect on

innovation and human resources), and change oriented (primary effect on innovation and adaptation to the environment). The theory states that effective leaders integrate behaviors in a way that is consistent with the situation and that the primary situational variables are the type of organization or unit and the amount of environmental volatility and uncertainty.

In general, situational models of leadership have received only moderate support from empirical studies. However, most of them have the advantage of providing detailed prescriptions for practitioners that are theoretically rooted in the emphasis of leaders' needs to interpret situations and to bring specific traits and acquired behaviors to these situations so as to exercise effective leadership.

3.2. Charismatic and Transformational Leadership

During the 1980s, the previous leadership models began to come under criticism due in large part to their dependence on the methodology used that gave rise to what is known as the "performance–cue" effect. This effect refers to the fact that groups with knowledge of their performance levels reported "objective" leadership behaviors that were consistent with performance levels (e.g., good leadership behavior was reported when performance was good, whereas poor performance led to less flattering ratings), regardless of the actual behavior of the leader. During these years, conventional leadership wisdom came under heavy attack, and this led to a "leadership gap" or "leadership crisis." Nevertheless, this gap was soon filled with a new kind of leadership whose characteristics were connected to one of the oldest traditions in leadership studies, that is, charisma. This Greek word meaning "special or divine gift" was already prominent in Weber's analysis of organizational authority in 1921. The charismatic leadership theory developed by Conger and Kanungo in 1988 relies on its effect on followers and society in general and stemmed from their radical vision of a new world, the exhibition of their own behavior for followers to imitate, and the confidence expressed in the followers' capacity to achieve challenging goals.

Therefore, this new perspective of charismatic leadership differs from the conventional line on leadership studies. To this point, this article has analyzed various aspects of leadership that have been grouped under the term "transactional leadership" because a type of transaction is established between a leader and followers in all of them. The members of the group or organization recognize the leader as such and accept his or her authority, but in turn the leader must provide the followers with valuable resources. As we have seen, these resources may occasionally refer to the cognitive skills and knowledge that enable objectives to be obtained, whereas other times they may be rewards such as pay increases and recognition of jobs well done. However, if one examines social reality and history itself, one finds that there is another type of leadership that produces important effects in followers in particular but also in society in general. This type of leadership has been labeled "transformational leadership" and refers basically to charismatic leadership.

The differences between transactional leadership and transformational leadership, as defended by Bass in 1985, are significant in terms of followers' commitment to challenging group or organizational goals. In transformational leadership, these goals are presented as a compelling vision and are expressed in ideological and inspirational language (e.g., "what we are able to achieve by challenging ourselves continuously"). In fact, transformational leadership is most likely to emerge when there is a group or organizational crisis or when followers are disenchanted with the current state of affairs. In other words, transformational leadership affects, first of all, followers' value systems and ambition levels. Its ultimate effect is that followers will appreciate projects and options that they would not appreciate unreservedly on their own. It transforms the value systems of followers in such a way that they start pursuing new ideals. Thus, the role of the transformational leader is to attempt to bring about followers' acceptance of a bold vision by inciting positive feelings toward desired behaviors. Change, then, comes about through an induced reevaluation of behavioral options by followers.

In contrast to transformational leadership, transactional leadership does not intervene in followers' value systems. Leaving these systems untouched, it tries to perceptually or materially model the contents of influence attempts to fit the existing values of followers. Followers become committed to supporting and implementing the proposals espoused by the leader because these proposals are intrinsically desirable and match their values and beliefs. The leader achieves this by clarifying or guaranteeing the profitability of requested efforts.

By the 1990s, issues of transformational leadership and charisma had become the dominant themes in leadership studies, and there was a proliferation of

models and perspectives on charismatic and transformational leadership. They emphasized, for instance, the connection between charisma and culture as well as between charisma and affective communication.

3.3. Individual Differences and Cognitive Processing Models

As theories on transformational leadership developed, new models in the domain of transactional leadership appeared, featuring a renewed interest in long since discredited trait-based models. Although trait approach research has been out of favor for a long time, recent work has once again suggested that stable individual differences in leader behavior do exist. Studies have become increasingly interested in the special endowments of leaders, and new models have focused on the intellectual and creative abilities of leaders. One good example of this development is the role motivation theory of organizational leadership that was introduced by Miner at the beginning of the 1970s and has been developing constantly ever since. The theory states that different leader attributes fit different types or roles in the organization and that leadership is a function of the larger organizational form within which it operates. Using a projective test, Miner described the type of motivational traits required for success in most leadership (management) positions in hierarchical organizations. Leader role requirements and the matching motive patterns are derived from the form of the organization. This person–organization fit model involves individual differences, strategy, organizational structure, and environment.

The individual differences research domain was strongly influenced by the information-processing concepts applied to the study of leadership perceptions. Cognitive research generally emphasizes the role of followers as well as leaders in leadership perception. Extensive research using an information-processing approach shows that individuals can be recognized as leaders based on their fit with leadership prototypes or that leadership can be inferred from favorable outcomes. Lord and colleagues' categorization theory and Mitchell and colleagues' attributional model represent that trend. The general idea is that leadership resides (at least in part) in the minds of the followers and that individual differences are important perceiver frameworks for assessing emergent leadership. This approach emphasized the role of individual differences in defining implicit theories or leader prototypes. The categorization theory assumes that relatively fixed prototypes, which are learned through extensive experience by followers, are the basis for leadership categorization. Followers appear to share (within a context) a set of general beliefs about the characteristics related to leadership (e.g., decisive, determined, intelligent) in diverse situations and tend to use their implicit theories and leader prototypes to decide whether or not an individual is judged to be an emergent leader. Specifically, if an individual is perceived to match followers' leader prototype, the person is likely to be viewed as an emergent leader and rated accordingly. From Lord and Maher's theoretical perspective, although leader qualities are clearly important for understanding leadership, followers' social–cognitive process is also important. Attributional notions are also introduced into the leadership cognitive approach, and the mutual influence of a superior's behavior and subordinates' behavior on one another is considered, inserting each person's causal attributions into the loop as mediational variables. As for future developments using the cognitive perspective, Lord and Emrich suggested in 2000 that one promising area covers the effect of self-concepts and self-regulatory processes such as self-efficacy. Although much of the research in the cognitive domain has focused on the relationship between a leader's behavior or traits and subordinates' satisfaction, behavior, and performance, increasing attention is being devoted to understanding the intervening mechanisms, such as values and self-concept, with which leaders influence followers.

As for conclusions on leadership research, Sternberg and Vroom provided a clear summary about theoretical leadership approaches in 2002: "The persons, the situation, and the interaction of the person with the situation . . . all matter for leadership effectiveness." Other recent reviews of leadership literature have called for increased attention on leadership context (e.g., Lowe and Gardner in 2000) as well as a greater focus on personal attributes, especially in the area of executive leadership (e.g., Hooijberg and Schneider in 2001). Similarly, Lord and Emrich suggested, in their review of cognitive leadership research in 2000, that "information processing research has established . . . that leadership factors reside not only in the minds of followers, but also in the minds of leaders, in leaders' behavior and attitudes, and in the social context in which leaders and their followers interact."

See Also the Following Articles

Authoritarianism ■ Leadership and Culture ■ Traits

Further Reading

Andriessen, E. J. H., & Drenth, P. J. D. (1998). Leadership: Theories and models. In P. J. D. Drenth, H. Thierry, & C. J. de Wolff (Eds.), *Handbook of work and organizational psychology* (2nd ed., Vol. 4, pp. 321–355). Hove, UK: Psychology Press.

Bass, B. M. (1985). *Leadership and performance beyond expectations*. New York: Free Press.

Bass, B. M. (1990). *Bass & Stogdill's handbook of leadership* (3rd ed.). New York: Harper & Row.

Conger, J. A., & Kanungo, R. N. (1988). *Charismatic leadership: The elusive factor in organizational effectiveness*. San Francisco: Jossey–Bass.

Fiedler, F. E. (1967). *A theory of leadership effectiveness*. New York: McGraw–Hill.

Fiedler, F. E., & Garcia, J. E. (1987). *New approaches to leadership: Cognitive resources and organizational performance*. New York: John Wiley.

French, J., & Raven, B. H. (1959). The bases of social power. In D. Cartwright (Ed.), *Studies of social power* (pp. 150–167). Ann Arbor, MI: Institute for Social Research.

Hardy, C., & Clegg, S. R. (1996). Some dare call it power. In S. R. Clegg, C. Hardy, & W. R. Nord (Eds.), *Handbook of organizational studies* (pp. 622–641). London: Sage.

Hersey, P., & Blanchard, K. H. (1977). *Management of organizational behavior: Utilizing human resources*. Englewood Cliffs, NJ: Prentice Hall.

Hollander, E. P. (1985). Leadership and power. In G. Lindzey, & E. Aronson (Eds.), *The handbook of social psychology* (3rd ed., pp. 485–537). New York: Random House.

Hooijberg, R., & Schneider, M. (2001). Behavioral complexity and social intelligence: How executive leaders use stakeholders to form a systems perspective. In S. Zaccaro, & R. Klimoski (Eds.), *The nature of organizational leadership* (pp. 104–131). San Francisco: Jossey–Bass.

House, R. J. (1971). A path–goal theory of leader effectiveness. *Administrative Science Quarterly*, *16*, 321–339.

House, R. J. (1988). Leadership research: Some forgotten, ignored, or overlooked findings. In J. G. Hunt, B. R. Baglia, H. P. Dachler, & C. A. Schriesheim (Eds.), *Emerging leadership vistas* (pp. 179–199). Lexington, MA: D. C. Heath.

Kelman, H. C. (1958). Compliance, identification, and internalization: Three processes of attitude change. *Journal of Conflict Resolution*, *2*, 51–60.

Kipnis, D. (1976). *The powerholders*. Chicago: University of Chicago Press.

Kipnis, D., Schmidt, S. M., & Wilkinson, I. (1980). Intraorganizational influence tactics: Explorations in getting one's way. *Journal of Applied Psychology*, *65*, 440–452.

Lewin, K. (1941). Analysis of the concepts whole, differentiation, and unity. *University of Iowa Studies in Child Welfare*, *18*, 226–261.

Lord, R. G., & Emrich, C. G. (2000). Thinking outside the box by looking inside the box: Extending the cognitive revolution in leadership research. *The Leadership Quarterly*, *11*, 551–580.

Lord, R. G., Foti, R. J., & De Vader, C. L. (1984). A test of leadership categorization theory: Internal structure, information processing, and leadership perceptions. *Organizational Behavior and Human Performance*, *34*, 343–378.

Lord, R. G., & Maher, K. J. (1991). *Leadership and information processing: Linking perceptions and performance*. Boston: Unwin Hyman.

Lowe, K. B., & Gardner, W. L. (2000). Ten years of The Leadership Quarterly: Contributions and challenges for the future. *The Leadership Quarterly*, *11*, 459–514.

Meindl, J. R. (1995). The romance of leadership as a follower-centric theory: A social constructionist approach. *The Leadership Quarterly*, *6*, 329–342.

Miner, J. B. (1993). *Role motivation theories*. New York: Routledge.

Miner, J. B. (2002). The role motivation theories of organizational leadership. In F. J. Yammarino, & B. J. Avolio (Eds.), *Transformational and charismatic leadership: The road ahead* (pp. 309–338). New York: Elsevier.

Mitchell, T. R., Green, S. G., & Wood, R. E. (1981). An attributional model of leadership and the poor performing subordinate: Development and validation. In L. L. Cummings, & B. M. Staw (Eds.), *Research in organizational behavior* (Vol. 3, pp. 197–234). Greenwich, CT: JAI.

Sternberg, R. J., & Vroom, V. (2002). Theoretical letters: The person versus the situation in leadership. *The Leadership Quarterly*, *13*, 301–323.

Vroom, V. H., & Jago, A. G. (1988). *The new leadership: Managing participation in organizations*. Englewood Cliffs, NJ: Prentice Hall.

Vroom, V. H., & Yetton, P. W. (1973). *Leadership and decision making*. Pittsburgh, PA: University of Pittsburgh Press.

Weber, M. (1921). The sociology of charismatic authority. In H. H. Gerth, & C. W. Mills (Trans. & Eds.), *From Max Weber: Essays in sociology* (pp. 245–252). New York: Oxford University Press.

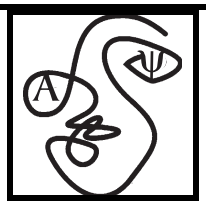
Yukl, G. A. (2002). *Leadership in organizations* (5th ed.). Upper Saddle River, NJ: Prentice Hall.

Yukl, G., (2003). Tridimensional leadership theory: A roadmap for flexible, adaptive leaders. In R. J. Burke, & C. Cooper (Eds.), *Leading in turbulent times*. London: Blackwell.

Yukl, G., & Falbe, C. M. (1990). Influence tactics in upward, downward, and lateral influence attempts. *Journal of Applied Psychology*, *75*, 132–140.

Yukl, G., & Tracey, B. (1992). Consequences of influence tactics used with subordinates, peers, and the boss. *Journal of Applied Psychology*, *77*, 525–535.

Zaccaro, S., & Klimoski, R. (Eds.). (2001). *The nature of organizational leadership*. San Francisco: Jossey–Bass.



Prejudice and Discrimination

Victoria M. Esses, Antoinette H. Semenya, and Monika Stelzl

University of Western Ontario, London, Ontario, Canada

1. Stereotypes, Prejudice, and Discrimination
2. Causes of Prejudice and Discrimination
3. The Experience of Being a Target of Prejudice and Discrimination
4. Psychological Strategies for Reducing Prejudice and Discrimination
5. Programs for Reducing Prejudice and Discrimination in Educational and Employment Settings

Further Reading

GLOSSARY

discrimination Negative behavior toward members of a social group based on group membership.

group identification The degree to which a person associates himself or herself with a group in terms of identity, shared interests, and outcomes.

group identity A person's view of, or sense of, himself or herself as a member of a social group.

illusory correlation Falsely perceiving a relationship between two variables where none exists.

in-group A group with which a person identifies or feels a sense of belonging.

minimal group paradigm Arbitrary assignment of individuals to different groups, with minimal information other than group membership provided.

out-group A group other than those to which a person belongs.

prejudice A negative prejudgment about or attitude toward members of a social group.

stereotypes Beliefs about the specific characteristics possessed by members of a social group.

Prejudice toward and discrimination against members of other groups can take a variety of forms and are caused by a number of factors. Their consequences for target group members are extensive and can potentially pervade all aspects of life. Several strategies for reducing prejudice and discrimination have been suggested, based on an understanding of the processes that serve to maintain negative intergroup relations. These strategies have led to the development of programs for use in educational and employment settings.

1. STEREOTYPES, PREJUDICE, AND DISCRIMINATION

People are judged and treated not only based on who they are as individuals, but based on the various groups to which they belong. For example, almost universally, members of the national group within a country are viewed more positively and receive better treatment than immigrant groups. Traditionally, men and women have been seen as possessing different characteristics, and women have frequently been viewed as inferior to men, with consequences for their treatment in society. Thus, relations among individual members of social groups can be strongly influenced by stereotypes, prejudice, and discrimination.

Although stereotypes, prejudice, and discrimination are conceptually related, they are not necessarily highly correlated. People may hold similar attitudes toward a variety of out-groups, whereas their stereotypes of these groups

©2004 Elsevier Inc.
All rights reserved.

may differ considerably. For example, although people may hold similarly negative overall evaluations of two groups, they may perceive one group as lazy and unintelligent and the other as conniving and cutthroat. In such cases, the stereotypes differ but the overall attitude toward the two groups is the same. Of importance, prejudice is not always related to discrimination. One major reason for this lack of association is that people do not necessarily act on their attitudes, particularly when such behavior may be prohibited by society. For example, even though an individual may hold a negative attitude toward members of a particular ethnic group, that individual might not be able to overtly express this attitude or blatantly discriminate against a member of the group due to societal constraints.

Some people claim that prejudice and discrimination are things of the past and that they no longer influence individuals in society. Although it is the case that overt, blatant prejudice has decreased considerably in recent years, it seems to have been replaced by a more subtle, covert prejudice. This subtle prejudice may be the result of unconscious negative emotions toward a group (of which prejudiced individuals are not even aware). It may be expressed in situations in which appropriate behavior is not clear-cut and where a variety of explanations for the negative behavior are available. For example, an individual might not hire an immigrant job candidate based on claims about the candidate's lack of local experience, rather than expressing negative attitudes toward the candidate's group.

Prejudice and discrimination may operate at different levels of society. At an individual level, a single person may hold prejudicial attitudes and discriminate against members of other groups. At an institutional level, prejudice and discrimination may operate through the practices and policies built into the institutional framework, which may prevent the full participation of members of certain groups. At a societal or cultural level, prejudice and discrimination may operate through the beliefs and values of the society that encourage and justify discrimination against particular groups. In all cases, the consequence is that certain groups receive negative outcomes and are not able to enjoy the benefits of full participation in society.

2. CAUSES OF PREJUDICE AND DISCRIMINATION

Prejudice and discrimination may be caused by a variety of factors. Chief among these are several factors that can be described as involving competition and threat. Realistic Group Conflict Theory and the more recent Instrument Model of Group Conflict propose that intergroup attitudes and behavior reflect group interests and are based, at least in part, on the nature of and the compatibility of group goals. When group goals are compatible, positive relations are likely to result. Conversely, when group goals are incompatible and groups are seen as competing for resources, conflict and negative intergroup relations are likely to result. Intergroup conflict, in turn, promotes hostility toward the competitive out-group. This hostility may be manifested in the expression of negative attitudes toward members of the other group and in attempts to prevent the other group from succeeding. For example, it has been demonstrated that when immigrants are seen as doing well and thriving in a difficult job market, members of the host population perceive increased competition with immigrants for resources and, therefore, are more likely to express negative attitudes toward immigrants and immigration. Relatedly, Social Dominance Theory suggests that threats to the in-group's power and dominance in society can lead to discrimination and willingness to aggress against less dominant out-groups. In such cases, prejudice may be driven by a fear of loss of social status or power rather than by a fear of loss of resources.

According to Social Identity Theory, for prejudice and discrimination to arise, it is not necessary to believe that a group is threatening tangible resources. Rather, the mere categorization of people into out-groups and an in-group stimulates a motivation to maintain or achieve a sense of positive group distinctiveness. As demonstrated in numerous studies using the minimal group paradigm, strategies for doing so include biases in evaluations that favor the in-group over out-groups and discrimination in resource allocations. In such cases, the motivation for a positive group identity may lead to processes that result in prejudice toward and discrimination against members of other groups.

Similarly, a number of recent theories have suggested that prejudice and discrimination may result from threats to values. For example, Symbolic Racism Theory proposes that prejudice against African Americans in the United States is partly attributable to a belief that African Americans threaten core American values, particularly those embodied in the Protestant work ethic (e.g., value of hard work, reward based on personal merit). Similarly, research on prejudice toward homosexuals has demonstrated that

perceived threats to religious and family values may form the basis of, and are used to justify, prejudicial attitudes.

Certain types of individuals are especially likely to perceive these types of threats and, thus, are particularly likely to hold negative attitudes toward other groups. Two personality variables that have been studied extensively in this context are right-wing authoritarianism and social dominance orientation. Right-wing authoritarianism involves submission to authority, belief in traditional values, and willingness to aggress against those who do not follow traditional ways. High right-wing authoritarians tend to be threatened by a variety of out-groups, particularly those that are seen as holding nontraditional values, and to hold negative attitudes toward these groups. In contrast, social dominance orientation involves a belief in inequality and preference for hierarchically structured social systems. Individuals who are high in social dominance orientation believe that unequal social outcomes and social hierarchies are appropriate, and they are especially likely to perceive threats to their group's status, power, and resources from other groups. As a result, they are especially likely to hold negative attitudes toward, and to discriminate against, other groups.

Group conflict and perceived threat, whether tangible or intangible, are not the only factors that can lead to prejudice and discrimination. There is a large body of research to suggest that people naturally categorize individuals into groups, particularly based on salient features such as sex, race, and age. Once people do so, a number of cognitive processes serve to promote stereotypes and prejudice toward members of other groups. For example, it has been demonstrated that the operation of an illusory correlation between a distinctive out-group (e.g., Asians) and a distinctive behavior (e.g., erratic driving) can result in an unfounded negative stereotype about members of the group (e.g., belief that Asians are bad drivers). Once formed, stereotypes are difficult to change due to selective attention to, interpretation of, and memory for stereotype-consistent information. In addition, once people categorize others into groups, they tend to exaggerate the similarities within groups and the differences between groups. Thus, members of an out-group are seen as all alike and as very different from members of the in-group.

In sum, the perception that a group is threatening the in-group's way of life and competing with the in-group for resources, power, status, and positive identity can lead to prejudice toward and discrimination against members of that group. Individuals who are high in right-wing authoritarianism and social dominance orientation are particularly likely to perceive such threats. Cognitive factors related to the categorization of individuals into groups can also promote and perpetuate negative stereotypes and prejudice.

3. THE EXPERIENCE OF BEING A TARGET OF PREJUDICE AND DISCRIMINATION

Individuals belonging to groups that are negatively stereotyped in society are potential targets of prejudice and discrimination. How do members of such groups experience and navigate the social world with these added concerns? Central to this experience is identity. Although considerable variation in identification exists, individuals belonging to devalued groups (e.g., ethnic minorities, physically handicapped individuals, homosexuals) tend to identify strongly with these groups. This heightened identity is the result of a combination of factors such as awareness of the devalued identity and availability of social support from group members. For many members of devalued groups, then a significant portion of self-image and identity is tied to that group membership.

Even before members of devalued groups become victims of prejudice and discrimination, the knowledge of their devalued group identity can have adverse consequences. One example of these consequences is evident in research on "stereotype threat." Stereotype threat occurs in situations in which members of negatively stereotyped groups perceive the possibility of confirming a negative group stereotype through their behavior. For example, the stereotype of African Americans as less intelligent than European Americans creates anxiety for them in academic domains. In these domains, performing poorly would serve to confirm the negative stereotype of low African American intelligence. When faced with the threat of confirming a negative stereotype, performance is impaired and, over time, disidentification with the academic domain can occur. As a classic demonstration of the stereotype threat effect, when a test is described as a measure of academic ability and race is made salient, the performance of African American students is impaired. That is, African American students perform worse

than European American students taking an identical test and also worse than African American students who are not told that the test is a measure of academic ability or for whom race is not made salient. Ironically, this is particularly the case for African American students who identify strongly with academic achievement. Extensive research has demonstrated the potential for such impaired performance related to the stereotypes of a variety of groups (e.g., women and math performance).

Because of the awareness of the potential for prejudice and discrimination, individuals belonging to devalued groups may exist in a state of attributional ambiguity. This term is used to describe a state in which members of devalued groups cannot make clear explanations for the causes of negative or positive events that occur in their lives. For example, if an African American individual receives poor service in a restaurant, it might not be clear whether this occurred due to the prejudice of a server or to the fact that the restaurant was busy at the time. This ambiguity can exist even when positive events occur. For example, if an African American student receives a high grade in a course, it might not be clear to the individual whether this grade was due to his or her ability or to the fact that the professor was concerned with appearing nonprejudiced.

This ambiguity surrounding the causes of positive and negative events can have positive or negative implications for the self. Making attributions to discrimination can protect self-esteem by allowing individuals to maintain positive self-beliefs in the face of failure in desired domains. However, making constant attributions to stable, external, and/or uncontrollable factors has been shown to be related to poor self-esteem, feelings of helplessness, and depression. Research shows that devalued group members at times will make attributions to bias, whereas in other situations they are reluctant to attribute events to prejudice. There is a tension that devalued group members must negotiate between making accurate assessments of prejudice and protecting perceptions of efficacy and control. Although the state of attributional ambiguity is described as pervasive, not all members of devalued groups will perceive and react to attributional ambiguity in the same way. Assessing the antecedents and consequences of attributional ambiguity is important in understanding the experience of being a target of prejudice and discrimination, as well as in promoting the overall well-being of these individuals.

There are other costs associated with making attributions to prejudice and discrimination, including social isolation, physical threat, financial loss, and emotional disturbance. Not surprisingly, individuals who belong to devalued groups are not always willing to perceive or admit prejudice and discrimination. One phenomenon that captures this is called the personal/group discrimination discrepancy. Researchers have found that individuals are willing to admit discrimination against their group in society, but are much less willing to admit that they personally have experienced discrimination. One factor that consistently affects the size of the personal/group discrimination discrepancy is individuals' identification with the group in question. Researchers have found that those higher in group identification are less likely to show the discrepancy; that is, they report more similar levels of personal and group discrimination than do those who are low in group identification. Perceiving or not perceiving discrimination has implications for numerous group processes such as identification and the willingness to engage in behavior that will help to improve the position of one's group in society.

In sum, the consequences of being a target of prejudice and discrimination are pervasive and begin even before observable negative behavior from others has occurred. However, individuals who belong to devalued groups are not passive recipients of biased treatment. They actively construct their world and experiences to protect and promote well-being and overall life success.

4. PSYCHOLOGICAL STRATEGIES FOR REDUCING PREJUDICE AND DISCRIMINATION

By being aware of the processes that lead to prejudice and discrimination, we can develop strategies for promoting more positive attitudes and behavior toward members of other groups. The contact hypothesis represents one of the most prominent approaches to prejudice reduction. In general, people are afraid of the unknown, which includes groups of people with whom they are fairly unfamiliar. However, the more knowledge about, understanding for, and experience with other groups that people obtain, the less negative and hostile people are likely to be. The underlying reasoning of the contact hypothesis is that if people are exposed to members of other groups in a variety of ways, they will become more positive toward members

of those groups and less likely to display prejudicial behavior. Nevertheless, it is not sufficient just to expose people to a minimal amount of information about another group or to live next door to someone who is of a different background. Several specific conditions must be met for the contact strategy to improve attitudes and behavior toward members of other groups.

Actual cooperation among members of the different groups is essential. In other words, people from the distinct groups should not be competing against each other for jobs, status, living space, money, and so forth. Cooperation between groups can be achieved by developing a common goal that can be achieved only if members of the different groups work together. Such cooperation will also help to develop a common identity that further decreases negative intergroup attitudes and discrimination. Equal status between the groups represents another important factor that underlies the contact strategy. In other words, it is critical that members of one group do not see themselves as superior to the members of the other group. Cooperative interdependence has been found to be helpful in superseding the presence of unequal status between two groups. In cooperative interdependence, each individual, regardless of group membership, is equally important for the task accomplishment.

Frequency, duration, and meaningfulness of interactions among the members of the different groups represent a set of factors that contribute to the effectiveness of the contact strategy. Specifically, the interaction among the members of the distinct groups should be frequent, long, and significant rather than occasional, short, and irrelevant. As a result, people will gain meaningful personal information about members of the other group. Social and institutional support is the final essential condition of the contact hypothesis. Contact between groups in the manner outlined here should be strongly supported by various levels of society such as the government, educational organizations, and employers. Overall, the contact approach provides a general framework consisting of several important factors that have been shown to contribute to decreasing prejudice and discrimination among members of different groups.

The second major approach to reducing prejudice and discrimination consists of categorization strategies. Two main strategies for restructuring people's mental representations of members of other groups have been suggested. Recategorization proposes the development of one common identity (i.e., us) rather than distinct in-group–out-group identities (i.e., us vs them). The existence of a shared identity will decrease the salience of the differences that often exist between two groups and will highlight the commonalities that the members of the two groups share. As a result, this focus on the shared aspects of a common identity will decrease the negative attitudes and discrimination between the two groups.

Decategorization offers a different approach to restructuring the perception of "us versus them." Rather than recategorizing separate identities by developing one shared identity, decategorizing signifies the elimination of group categorization. In other words, the removal of group-based typing will lead to the individualization of the members of the other group. As a result, they will be perceived as unique persons without the negative stereotypes that are traditionally associated with that specific group. This perception of people as unique individuals will further decrease the prejudice and discrimination that were brought about by intergroup differentiation.

It is also important to address the role of individual differences in the effectiveness of strategies for the reduction of prejudice and discrimination. For example, to decrease prejudice and discrimination, it is important to develop strategies that will counter right-wing authoritarian tendencies. One strategy that has received support is the personal value confrontation technique. With this technique, the discrepancy between people's ratings of the importance of freedom versus equality and their simultaneous support for discriminatory treatment of a minority group is highlighted. When the attention of high right-wing authoritarians is drawn to this incongruity, they tend to show improved attitudes toward the minority group.

In sum, based on knowledge of the processes that underlie prejudice and discrimination, strategies to counteract these influences can be suggested.

5. PROGRAMS FOR REDUCING PREJUDICE AND DISCRIMINATION IN EDUCATIONAL AND EMPLOYMENT SETTINGS

Several programs for reducing prejudice and discrimination in educational and employment settings have been implemented. As shown in Table I, one way of thinking about these programs is to categorize them

TABLE I
Programs for Reducing Prejudice and Discrimination in Educational and Employment Settings

Enlightenment-based programs
- *Multicultural education:* Provide information on the history and cultural practices of a wide variety of groups and emphasize standards of acceptance and inclusiveness
- *Cultural diversity training:* Attempt to promote value of group differences and understanding of how behavior is influenced by people's culture and background
- *Moral development training:* Emphasize standards of behavior and values of social justice, fairness, and equality

Interaction-based programs
- *Intergroup dialogues:* Allow members of different groups to address areas of contention through discussion and recognition of different perspectives
- *Cooperative learning and cooperative work groups:* Encourage cooperative interdependence among members of different groups working toward a common goal

as enlightenment-based programs or interaction-based programs.

The enlightenment-based programs generally involve increasing the knowledge that people have of other groups and altering people's perceptions of their relationships with others. These programs focus on changing the representation of members of other groups (e.g., reducing stereotypes, personalizing others, developing more inclusive representations) and on highlighting values relevant to tolerance and equality. For example, cultural diversity training programs are brought into industrial settings to teach managers and employees to value group differences and to develop increased understanding of other groups. Through this process, it is hoped that individuals will become committed to identifying and preventing the operation of stereotypes and prejudice in the workplace.

Interaction-based programs focus on bringing members of different groups together to provide direct interaction among members of the groups. They exert their influence through processes associated with intergroup contact. For example, cooperative work groups bring members of different groups together to work on a common task and emphasize cooperative interdependence whereby each person is dependent on the others for successful completion of the task. In the process, it is hoped that members of the different groups will come to appreciate the contributions of all others, come to see each other as members of the same "team," and learn about the individual characteristics of others so that they are personalized.

Despite the plethora of programs to reduce prejudice and discrimination in educational and employment settings, the effectiveness of many of these programs has not been fully evaluated. Indeed, assessment of effectiveness depends critically on the specific outcomes desired. For example, is the goal to reduce stereotypes, change group representations, or reduce tension among members of different groups? Depending on the context and on the rationale for implementing a program in a new setting, one or another of these goals may be paramount. As a result, in choosing which program to introduce in an educational or employment setting, one must first identify the specific goals desired. Once an appropriate program is implemented, a continuing process of evaluation of effectiveness can be put into place.

See Also the Following Articles

Employment Discrimination ■ Homosexuality ■ Interpersonal Conflict ■ Stereotypes

Further Reading

Brewer, M. B., & Brown, R. (1995). Intergroup relations. In D. T. Gilbert, & S. T. Fiske (Eds.), *The handbook of social psychology* (4th ed., Vol. 2, pp. 554–594). New York: McGraw–Hill.

Brown, R. (1995). *Prejudice: Its social psychology*. Oxford, UK: Blackwell.

Crocker, J., Voelkl, K., Testa, M., & Major, B. (2000). Social stigma: The affective consequences of attributional ambiguity. In C. Stangor (Ed.), *Stereotypes and prejudice: Essential readings* (pp. 353–368). Philadelphia: Psychology Press.

Esses, V. M., Jackson, L. M., & Armstrong, T. L. (1998). Intergroup competition and attitudes toward immigrants and immigration: An instrumental model of group conflict. *Journal of Social Issues*, *54*, 699–724.

Fiske, S. T. (1995). Stereotyping, prejudice, and discrimination. In D. T. Gilbert, & S. T. Fiske (Eds.), *The handbook of social psychology* (4th ed., Vol. 2, pp. 357–414). New York: McGraw–Hill.

Gaertner, S. L., & Dovidio, J. F. (2000). *Reducing intergroup bias: The common ingroup identity model*. Philadelphia: Psychology Press.

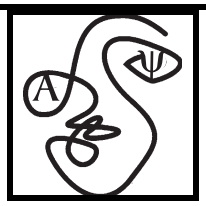
Oskamp, S. (Ed.). (2000). *Reducing prejudice and discrimination*. Mahwah, NJ: Lawrence Erlbaum.

Sidanius, J., & Pratto, F. (1999). *Social dominance: An intergroup theory of social hierarchy and oppression*. New York: Cambridge University Press.

Steele, C. M., Spencer, S. J., & Aronson, J. (2002). Contending with group image: The psychology of stereotype and social identity threat. In M. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 34, pp. 379–440). San Diego: Academic Press.

Stephan, W. G., & Stephan, C. W. (2001). *Improving intergroup relations*. Thousand Oaks, CA: Sage.

Swim, J. K., & Stangor, C. (Eds.). (1998). *Prejudice: The target's perspective*. San Diego: Academic Press.



Privacy

Carol M. Werner, Barbara B. Brown, and Irwin Altman
University of Utah, Salt Lake City, Utah, USA

1. Background and Definition
2. Importance of Privacy Regulation
3. Loneliness, Crowding, and Crime: Consequences of Adequate and Inadequate Privacy Regulation
4. Technology and Privacy Regulation
5. Summary and Conclusions
Further Reading

GLOSSARY

crowding Perception that there are too many people in a situation.

privacy as a cultural universal The notion that all societies have developed mechanisms for regulating openness and closedness.

privacy mechanism Behaviors used to change the amount of openness or closedness in an interaction.

privacy regulation Selective control of access to the self or to one's group, that is, control over the degree of openness or closedness.

Privacy regulation refers to how individuals and groups control interactions with others, especially about confidential or intimate information or with respect to physical contact or access. All societies have developed mechanisms for regulating privacy, and effective use of these mechanisms is related to individual and group viability.

1. BACKGROUND AND DEFINITION

1.1. Definition

According to Altman in 1975, privacy regulation is defined as the "selective control of access to the self or to one's group." Although standard dictionary definitions equate privacy with seclusion and withdrawal, psychologists study privacy regulation processes as a dialectic wherein people ebb and flow between seeking and avoiding social contact. There are many applications of privacy theory, including privacy policies in hospitals and other businesses, privacy policies on Web sites, security screening, and planning for privacy regulation in the design and use of home and community settings. All of these applications must accommodate people's needs to be open as well as closed or else they will be inadequate. Creating information management systems or designing settings that only allow people to close themselves off makes it difficult, if not impossible, for people to share information and/or connect with others, whereas systems and designs that encourage openness, sharing, and accessibility run the risk of preventing people from guarding their own secrets and/or screening out unwanted information from others.

"Selective control" means that people aim to manage whether, when, and with whom to be open or closed. They actively negotiate how accessible they want to be about themselves and their personal information. Decisions about openness/closedness, as well as

©2004 Elsevier Inc.
All rights reserved.

abilities to control degrees of openness, vary with time and circumstances. For example, among friends and acquaintances, there are some people to whom an individual is almost always available (regardless of his or her other desires, obligations, and activities), some people the individual talks to unless he or she is busy, and other people the individual avoids most of the time. With respect to strangers, the individual may ignore them completely or may be open in civil and polite ways. In some situations, the individual may eagerly seek interactions with strangers and view them as opportunities to form new friendships, interact across cultures, and/or create a more civilized community (e.g., pleasantries in a park or at sporting events). At the same time, it is important for the individual to maintain reserve with strangers until safety is ensured. The general principle is that people constantly and selectively choose how open or closed they desire to be. People often act in ways that help to achieve their goals, creating blends of openness and closedness for short or long periods of time.

"Access" refers to a two-way process of both receiving information from others and giving information to others. People monitor both what comes in and what is given out, and they decide whether to allow the information to pass. Information includes physical stimuli such as sounds, odors, visual information, and even the presence of someone at the door or on the telephone. Other information that people aim to control is economic and financial such as social security numbers, credit card numbers, and banking information. Information can also be more psychological in nature such as intimate and confidential information about people's criminal or sexual histories, health problems, grades, and incomes. People try to maintain control over what they learn about others and what others learn about them. The importance of controlling information varies with the nature of the information, individual preferences, time, and circumstances, and people put more or less effort into privacy regulation for different reasons.

"Self or group" is included in the definition to emphasize that sometimes individuals are the focus of analysis and sometimes dyads and groups are the focus. That is, individuals regulate interaction with other individuals or groups, and dyads and groups often act in coordination to regulate interaction with outsiders. For example, privacy regulation within a marriage involves regulation between the spouses, between the spouses and their respective in-laws, and between the spouses and their friends or outsiders.

In sum, according to Altman,

> Privacy is an interpersonal boundary-control process, which paces and regulates interaction with others. Privacy regulation by persons and groups is somewhat like the shifting permeability of a cell membrane. Sometimes the person or group is receptive to outside inputs, and sometimes the person or group closes off contact with the outside environment.

Thus, privacy regulation is dynamic, and people constantly monitor their needs and opportunities and then decide what levels of openness and closedness to seek or accept.

1.2. Privacy Mechanisms

People use many "mechanisms" or strategies for achieving their desired levels of openness and closedness. Some mechanisms are verbal or paraverbal (e.g., intonations of words), whereas others are nonverbal (e.g., interpersonal distance, eye contact). Many mechanisms involve the physical environment; examples include opening or closing doors, choosing more sociable or more secluded settings, and creating territories by claiming and defining places and their boundaries. Many mechanisms are technological; examples include answering machines and caller ID that allow an individual to selectively answer or ignore telephone calls, filters on e-mail that allow blacklisting of "spam," and home-centered filters on television channels and Web sites to control what information children can access. Mechanisms that facilitate openness include instant messaging, videoconferencing, and the ready availability of cell phones.

Individuals can use privacy mechanisms alone or in combination to seek out or exclude others and information about others. With respect to the intimacy level in an interaction, sometimes a subtle facial expression is sufficient and sometimes a facial expression, physical distance, and changing the topic all are needed to achieve the desired level of intimacy. Similarly, with respect to physical contact, sometimes a single mechanism is sufficient and sometimes additional mechanisms are needed. Some individuals who are very skilled at privacy regulation can achieve contact with a telephone call and/or can achieve needed solitude without offending others, thereby leaving open future opportunities for interaction. As an example of the latter, consider the difference between "Sorry, I'm busy right now" and "Sorry, I'm busy right now but I'll call you when I'm free." The second version achieves

short-term solitude while maintaining opportunities for future openness.

2. IMPORTANCE OF PRIVACY REGULATION

2.1. Importance

Privacy regulation serves critical individual and interpersonal functions. Individuals need to be able to regulate how much stimulation they receive and how much information others have about them. They need to control their physical environment and keep control over resources such as food, time, and space. Individuals and groups that can be easily intruded on, stolen from, forced to disclose secrets, and/or violated in other ways can lose a sense of self. Similarly, individuals and groups that cannot connect with others for personal or professional goals are isolated and unable to participate in interpersonal activities that are essential for functioning effectively in society. Individual and group viability is an important consequence of effective privacy regulation.

2.2. Privacy Regulation as a ''Cultural Universal''

Because of its importance to effective functioning, privacy regulation is acknowledged as a "cultural universal," meaning that every society has developed mechanisms so that individuals and groups can regulate interactions with others. Particular privacy mechanisms vary from one group to another, and use of privacy regulation mechanisms is usually guided by cultural norms, customs, and styles of behavior. Examination of societies around the globe shows that they vary in their preference for being relatively open or closed, but even the most open societies have mechanisms for closing themselves to others and even the most closed ones have ways of opening themselves up to others.

Most societies treat homes as places where individuals can withdraw, knowing that they can develop selected relationships and will rarely experience unwanted visitors. Visitors are expected to call ahead or announce themselves in other ways and then wait to be invited inside. In some situations, residents signal their availability to guests. For example, in traditional Egyptian Nubian homes, decorative plates are mounted on the homes' exteriors but are removed when the families are in mourning. The presence or absence of the plates signals to visitors whether they would be welcome. The Navajo in the southwestern United States have a custom that gives control over visitors to residents: Visitors do not knock, but they make themselves visible and wait for residents to invite them in. In some societies or for some groups of people, it is acceptable to use "keep out" mechanisms if people so choose, and no offense is taken when people leave the room, go on extended trips, or simply ignore visitors. In other groups, these and other "keep out" mechanisms are acceptable if people are busy working but are less acceptable if people are involved in other activities such as entertainment and loafing. The statement that someone "needs space" affirms the idea that people can legitimately vary between degrees of openness and closedness.

There are many privacy regulation mechanisms for the control of confidential information, and their use depends on many factors. Interpersonal distance controls access to some personal information (e.g., odors such as alcohol on the breath, physical details about age such as gray hair and wrinkles). Controlling verbal information is quite common as people choose to disclose or not disclose confidential information. In some groups or for some people, not only refusing to disclose but also actively lying is acceptable as a privacy strategy. Although people may use evasion or lying to control information, in many societies (e.g., the Javanese in Indonesia) lying is rarely needed because confidentiality is respected and it is considered rude to inquire too much about other people's business.

In 2002, Petronio suggested that privacy rules and mechanisms often change and evolve so that they work most effectively for each group. She noted that many groups negotiate both general and specific privacy regulation rules. An example would be a group that shares confidential information and understands that the information should not be disclosed to the general public but negotiates specific details such as whether spouses or particular friends could be told some part of the story. Thus, the group co-owns the information, agrees to cooperate in protecting it, but negotiates particular boundaries around different parts of the information.

As computers make sharing of confidential information easier, mechanisms are being developed to control who has access and for what reasons. Privacy policies on Web sites promise that credit card information is protected from theft, that one can choose whether to be contacted, and that one's contact information is not shared without permission. An important difference

between U.S. and European Union rules lies in the "opt-in" versus "opt-out" wording of these policies. In the United States, a company assumes that it can use or even share contact information for promotional purposes unless consumers explicitly notify the company to keep their identities confidential. In contrast, in the European Union, the opposite is assumed, and consumers must notify the company of their willingness to share identity information.

3. LONELINESS, CROWDING, AND CRIME: CONSEQUENCES OF ADEQUATE AND INADEQUATE PRIVACY REGULATION

3.1. Dynamics of Privacy Regulation

As illustrated in Fig. 1, privacy regulation is a dynamic and ongoing process. People develop a desired level of contact with others, compare that desired level with their actually achieved level, and activate privacy regulation mechanisms (if needed and possible) designed to reach their desired levels of contact. When people desire more contact than they are experiencing, they feel lonely and select mechanisms that will bring them into contact with others. When people desire less contact than they have achieved, they describe themselves as crowded or overwhelmed and select mechanisms that will reduce their levels of contact.

Judgments about crowding and loneliness are very subjective and can be based on interaction partners, topics of conversation, and whatever alternative activities might be available. An individual can feel lonely in a crowd of people if the particular people the individual wants to be with are absent. The process can be very fluid. An individual can join a group and feel that his or her social contact is optimal, spend some time with that group but then begin to feel crowded or lonely, and then move to a different group and feel that social contact is optimal again. Problems with privacy regulation can occur when individuals or groups lack the behavioral and interpersonal skills needed to seek out or avoid contact with others or when the physical or social environment makes privacy regulation difficult.

3.2. Viability

The ability to manage privacy is related to viability, that is, the ability to thrive, achieve goals, grow, and adapt to a changing environment. Research suggests that individuals and groups that do a better job of regulating interactions are more viable. Measures that tap viability have varied and include psychological variables such as satisfaction with the environment, satisfaction with social relations, improved task performance, greater task persistence, less stress, and other evidence that people are successful in achieving their short- and long-term goals. Several lines of research

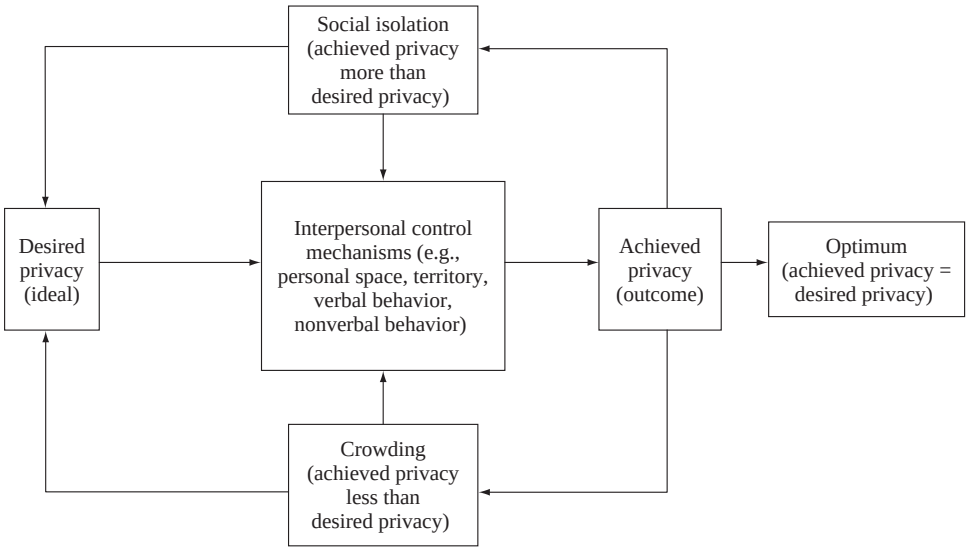

Social isolation (achieved privacy more than desired privacy)
Desired privacy (ideal)
Interpersonal control mechanisms (e.g., personal space, territory, verbal behavior, nonverbal behavior)
Achieved privacy (outcome)
Optimum (achieved privacy = desired privacy)
Crowding (achieved privacy less than desired privacy)

FIGURE 1 The process of privacy regulation.

have addressed relationships between privacy regulation and indexes of viability. These include loneliness, crowding, family and community relations, burglary and crime, and interaction stress in offices, prisons, hospitals, and other public settings. Typically, a combination of intrapersonal and sociophysical factors contribute to viability. For example, in one study, when pairs of naval recruits began an isolation mission by clearly establishing and maintaining separate territories that gave order to and allowed control over daily activities, they were more likely to complete the mission successfully. In contrast, pairs who did not create separate spaces did not perform as well on daily assignments and were more likely to abort the mission before the completion date. Thus, viability was related to successful management of privacy.

Another study asked dormitory residents what strategies they used for privacy regulation, that is, what they did when they wanted to make contact with others and what they did when they wanted to avoid contact. These strategies were related to one index of viability: whether the students chose to return to school the following year (excluding academic expulsions). Again, viability was related to effective strategy use. Students who remained in college were more likely to use both opening and closing mechanisms and to state that the mechanisms were successful. Students who dropped out reported using fewer of the successful mechanisms and also reported using other mechanisms that did not help them to achieve their desired levels of contact.

3.3. Loneliness and Privacy Regulation

Loneliness is usually considered to be a problem stemming from a lack of individual skills, but solutions include both skills training and restructuring of the physical environment. Clinical and counseling psychologists have developed training programs designed to help people learn how to make and maintain friendships, and some of these programs have been quite successful. Equally important is research suggesting that physical and social environments play a role in supporting or discouraging contact. For example, one university study found that foreign students could regulate interactions with family members inside their apartments but had difficulty in making friends with residents in the courtyard. One explanation is that the foreign students had difficulty because the customs and strategies for making contact in the United States were different from those in their countries of origin; for example, the Asian students did not feel comfortable in making casual "small talk," whereas that was how U.S. students tended to initiate interactions. The housing director began to provide additional settings in which U.S. and foreign students could interact. Another study found differences in friendship patterns in inner-city public housing developments to be related to the physical environment. These researchers found that having green spaces close to residents' apartments was related to increased informal contact among neighbors, larger social networks among adults, and higher rates of child–adult contact and interaction in the outdoors. In essence, residents were drawn to green spaces, and this increased proximity led to informal contacts that then led to deeper friendships.

The importance of the physical environment in providing opportunities for friendship formation is well known from research on proximity in housing developments as well as from research on public gardens as a context for friendship formation. A growing trend in community development is to use design features that invite people to spend time in community settings. These include streets that invite walking (e.g., storefronts at the sidewalk instead of behind parking lots, safe tree-shaded streets with attractive architecture, few cars or low speed limits, things to look at along the way, proximity to shopping and other services) and the availability of parks and green spaces where children and adults can congregate. These studies combine to suggest that coupling social skills training with supportive physical and social environments is all important for helping people to avoid loneliness.

3.4. Crowding and Privacy Regulation

One of the most well-researched areas in privacy regulation is crowding. Unlike being "cramped," where one person has too little space to accomplish tasks, crowding always involves more people than one desires or too little space for those people in a setting. Crowding occurs when an individual or group feels too accessible to others. For example, the phrase "two's company, three's a crowd" captures the idea that it is not the sheer number of people but rather one's goals in an interaction that determines whether one experiences crowding. Similarly, being part of a crowd at a football game not only is not crowding but actually part of

the fun of being a fan is the presence of numerous others.

Theories and research on crowding have focused on "control" and the lack of control that accompanies the experience of crowding. Loss of control is an important construct because it connects the work on privacy regulation—and the theory's central idea of selective control—with research on stress and stressors. Stressful events are often characterized by a loss or lack of perceived control. Sometimes, people experience crowding stress because they cannot control the amount or type of stimulation they receive from others such as crowd noises that overwhelm them. Related to this is crowding stress that occurs from uncontrolled interaction with others such as too many people at a party and too many neighbors in a housing complex. Sometimes, stress occurs because people interfere with or frustrate each other's goal attainment, as might occur when too many people try to get on a subway or freeway, when a third person thwarts dyadic intimacy, or when there is no space to sit at the park or on the beach. Another common basis of crowding stress is resource based and occurs when there are too many people for the available food, water, and space. All of these bases of crowding are subtle variations on the idea that other people can undermine people's sense of control over their own goals as well as over their interactions with others.

Studies conducted among college and university dormitory residents revealed that crowding was related to negative effects on friendship formation, lower satisfaction with and commitment to their academic institution, and lower grades. Housing three people in a room resulted in unexpected social dynamics. Male students tended to withdraw from the situation and leave the room to study at the library or spend time with friends. Female students tended to form "two against one" coalitions that left one student feeling alone and rejected. These problems were reversed when students were housed two to a room.

Social dynamics are also a problem in large dormitories (i.e., 32–40 residents) compared with both dormitories providing differentiated spaces and smaller dormitories (i.e., 6–20 residents). Sharing a dormitory with large numbers of residents makes it difficult to learn individual names and differentiate one student from another. It also makes it difficult to avoid social contacts when walking down the hall or using the shower. Initially, students in large residence halls may try to assert control and regulate their interactions, but eventually they stop trying as much to do so. In one experiment, students in large dormitories reported a variety of social and educational consequences. They spent less time socializing, began to ignore other residents, complained about greater difficulty in controlling interactions, and said that they had more unwanted interactions. They also reported evidence of learned helplessness such as putting less effort into trying to control interactions with others and finding quiet places to study.

Research also shows that although people can tolerate high density for short periods of time, sustained high density for several weeks is stressful by itself and can exaggerate other stressors, including the stress of daily hassles. To add to these problems, friendships often deteriorate under the stress of high-density living, and people lose opportunities for social support that would ordinarily help to buffer them against stress.

In addition to helping to understand relationships between density and stress, these studies have been valuable in showing design features that give more control to residents and reduce crowding stress. For example, dormitories are now designed as suites or apartments to create smaller subgroups of residents. Breaking up long hallways into separate areas and providing kitchens, bathrooms, and social areas for each group gives residents manageable numbers of interactions, gives more control over who uses their spaces, reduces competition for resources, and restores order to interactions. Suite or apartment designs house six or more students together, with central common areas surrounded by separate sleeping areas. The common living room and bathroom encourage residents to function as a group, whereas the separate sleeping rooms each house two students, allowing them to withdraw from the group as needed but also allowing them to form closer relationships with a smaller number of residents. In theory, each cluster or pod then becomes a viable social unit that can interact with other similar clusters in the larger dormitory building.

Research has also examined how to optimize functioning in homes and apartments, regardless of whether residents are related, involved, or living together as roommates. Some groups develop effective rules that allow individuals to regulate interactions with others. People can use open and closed doors and verbal messages (e.g., "I need to study now") to signal their accessibility and inaccessibility. In many cases, design features can support such regulation. People can withdraw to their bedrooms, can study, or can use the bathroom if they desire to be alone.

Research indicates that homes and apartments with highly differentiated spaces (i.e., more small rooms with separate access via hallways) give residents more control over their interactions (i.e., allow them to decide whether to be open or closed to others) and reduce crowding stress. The ability to avoid contact must be complemented by opportunities for making contact. In most apartments and homes, people can gather in common areas, such as living rooms and family rooms, when interaction is desired. Often, people are attracted to rooms with activity opportunities such as television sets, games, and stereos. These areas become "behavioral focal points," that is, spaces whose design, location, and function invite people to gather together.

Another design feature that supports privacy regulation is the "privacy gradient," that is, the idea that places in homes vary in their physical accessibility to outsiders. This physical accessibility should parallel residents' desired accessibility. Some rooms and spaces are close to the entrance, whereas others are deeper in the home's interior and so are less accessible to outsiders. The front yard, main entrance, and entryway all are fairly accessible to anyone who desires contact. Living rooms, dining rooms, and guest bathrooms tend to be formal areas where strangers and visitors might be invited. Places such as the kitchen, family room, bedrooms, and family bathrooms are typically used exclusively by residents. Managing interactions may be facilitated when the physical layout of the home or apartment puts public areas closer to the entrance and private areas far from the entrance and less accessible to visitors.

3.5. Privacy Regulation in Other Settings

Research on residential crowding has been paralleled in prisons, homeless shelters, hospitals, and similar settings, and the results have generally been similar. Living in undifferentiated spaces with numerous others is stressful. Aggression and crime can be high, and residents feel vulnerable and unable to control interactions. For example, in dormitory-style prisons, there were high aggression rates as well as high stress-related illness rates. Redesigning these spaces with simple barriers that gave each inmate a separate area around a bed reduced stress levels considerably. Consistent with the idea that privacy regulation involves openness and closedness, the prisons did not completely enclose each prisoner. Walls were low enough and entryways had no doors, so that guards could ensure safety for each prisoner and monitor for illegal activities.

3.6. Crime and Defensible Design

Public housing in the United States has become notorious for its physical unattractiveness, lack of upkeep by residents and staff, and high crime rates that include drug abuse, minor and major property crimes, gang warfare, and even murders and drive-by shootings. For more than 50 years, social scientists have tried to understand the bases of this disorder so as to correct it. A central issue is the extent to which the housing developments themselves undermine positive social relationships. Comparisons of different configurations indicate that developments with more "defensible designs" yielded more viable communities. Design elements allow residents to differentiate legitimate users from outsiders, encourage residents to use outdoor spaces and become protective of their own territory, and allow residents to use passive surveillance to monitor legal and illegal activities. In this way, defensible design can be viewed as a privacy regulation mechanism. There is no simple panacea to problems associated with many housing developments, but some essential design features include limiting the numbers of residents so that they can recognize one another and distinguish residents from outsiders (this is often accomplished with relatively low numbers of floors and limited numbers of apartments on each floor and with separate entrances and separate elevators), windows that allow surveillance of the outside area, outside gathering areas that provide activities and seating for multiple age groups, attractive settings to increase pride in place, and clear differentiation of the private setting from the public street.

Most of the research on defensible design has been done in public housing developments. These ideas have been replicated in middle-class suburban neighborhoods, where homes and neighborhoods with more defensible space are less likely to be burglarized. Neighborhoods are less likely to attract burglars when they appear to have stronger territorial claims, that is, when the homes are personalized and the homes and neighborhood convey the impression that residents feel a sense of ownership over the space and would react to the presence of suspicious strangers. The roads appear to be less public (i.e., fewer traffic signs of all types), sites allow surveillance from one home to another, residents appear to spend time on the block, and intruders are often challenged by residents or reported

to the police. Sense of territorial control is also apparent at each home site. The home and yard are personalized with names and house numbers, and there is often evidence that someone is at home (e.g., toys and tools are in the yard, mail and garbage are picked up). Surveillance is clearly possible because shrubs do not obscure the entrance or the windows, and homes are sited so that neighbors can observe each other's homes. When burglars are asked to indicate which homes they are not likely to target, they usually indicate the homes where they would expect the resident or a neighbor to challenge them if they tried to enter.

In sum, the physical environment serves important privacy regulation functions. Environments that provide clear differentiation between public and private areas are less likely to be intruded. Neighborhoods and housing developments whose designs support neighborly friendships also support creation of neighborhood territories and defensible spaces. These spaces also contribute to enhanced control over who can enter and who cannot as well as to what people can do in these spaces.

4. TECHNOLOGY AND PRIVACY REGULATION

4.1. Communication Technology and Privacy

As people increasingly use technology to connect them to the world, people are also developing technological mechanisms that allow them to control their accessibility and inaccessibility to others. People use cell phones, telephones, fax machines, e-mail, and the World Wide Web 24 hours a day to contact friends and business associates. They can meet new people, learn new ideas and philosophies, try on new personas, and even fall in love using Web-based exchanges. They can also send pictures through their telephones and the Web. Television is expected to allow Web access that will expand the ability to select programs, movies, sports events, and so on. Sophisticated cameras provide surveillance to increase safety in parking garages, on university campuses, and elsewhere, and cameras are used to cite motorists for speeding and other violations.

Each of these opportunities comes laced with issues essentially involving tensions between openness and closedness. How and when do people control their accessibility to others? How can people trust the honesty of others they meet online? If people have cell phones, should they be available 24 hours a day? Does owning a telephone and having telephone service mean that telephone solicitors can contact people, or can people decide who can call them and who cannot? Do Web and television access mean that children can view material that is intended for adults? How can children be protected from Web-based crime, including adults who use chat rooms as a base for contacting and harming children? When does surveillance become intrusive?

Extensive discussions of these and related issues have begun to lead to a variety of solutions. This article addresses proposals in the United States, although other countries are dealing with similar issues. Efforts are ongoing to find ways of modifying and improving these control systems so that they provide opportunities for openness while protecting people from serious harm as well as minor inconveniences. Just as different but effective privacy mechanisms have evolved in different cultures, privacy mechanisms are continuing to evolve as new technologies open the world to people. Some strategies that allow selective access already exist such as caller ID on telephones, unlisted telephone numbers, listed numbers that omit home addresses, and Web search engines that screen children from pornography but allow adults wider access.

At the same time, marketers selectively target particular consumers using information gathered on the Web, grocery store "preferred customer" spending patterns, and credit card files and other sources. Advertisers seek to use this personal information to target marketing to the most appropriate and receptive consumers. They emphasize the importance of free and open access to consumers as a way of maintaining a vital economy. Knowing individuals' family sizes and configurations, buying habits, business and recreation preferences, annual incomes, and other information allows companies to target consumers with promotional materials. Some consumers enjoy receiving this material, regardless of whether it comes through the Internet, telephone, or mail, whereas others consider these contacts to be intrusions and seeking ways of preventing such marketing. Privacy regulation strategies include state and federal legislation as well as commercially available screening systems and personal controls.

Faxes and telephone calls are regulated by the Telephone Consumer Protection Act of 1991. When fax machines first became widely available, marketers sent unsolicited promotional materials. Because consumers could show evidence of harm—the recipient had to pay for paper and ink, and incoming offers tied up their fax machines—federal rules supported

recipients over senders. Faxes could not be sent without permission or an existing relationship (e.g., organizations could fax conference announcements to members). A more restrictive rule, requiring recipients' written permission before faxes can be sent, was under review at the time of this writing.

Similarly, telephone solicitors created sufficient problems (e.g., calling at dinnertime; taking financial advantage of people) that federal policies and supportive legislation have been put into place to allow people to put their names and telephone numbers on "do not call" lists, with no charge to consumers but with hefty penalties to marketers for violations. The policy was challenged in court and was on appeal at the time of this writing. This list is for interstate commerce, and each state must enact its own "do not call" list to regulate local marketing calls. Some states want consumers and marketers to share the cost of this service, indicating their belief that telephone solicitors have a right to make calls and that consumers should pay to assert their right not to be called.

E-mailed solicitations (derisively called "spam") have been more difficult to stop, in part because there is less apparent harm to consumers and marketers claim that unsolicited e-mails could be seen as a necessary nuisance. Also, many senders immediately change their e-mail addresses or otherwise conceal their identities, making it difficult to screen future mailings or locate the senders if the behavior is declared illegal. Evidence of harm has been claimed by Web administrators whose server equipment must handle large numbers of unsolicited and unwanted e-mails daily. These administrators are pressuring the federal government to stop spam.

4.2. Health Care Privacy

The appropriate blending of openness and closedness is also a central issue in health care privacy. Patients may wish to be accessible by sharing test results among several doctors or by receiving information about new medications and medical procedures for their particular conditions, they may need an efficient way in which to get billing information to their insurance company, they may wish to receive test results on the telephone, and so on. Therefore, delimited openness of medical information can be beneficial. On the other hand, patients may wish to disseminate only information about particular medical conditions, may be uninterested in hearing from marketers, may want to withhold some information from insurance companies (even to the point of paying privately for some tests so that insurance companies will not perceive them to be risky clients), may want to receive test results only on a confidential telephone line (rather than having a coworker take a message), and so on. A major concern has been that patients' fears about control over information might make them unwilling to be open and honest with their doctors, thereby putting their health at risk. Patients might not disclose illegal or embarrassing behaviors, past health problems, health problems of relatives, and other private concerns, lest this information be widely disseminated without their explicit permission. People may opt out of genetic tests that can yield important information because they fear that employers might get the information and use it for insurance coverage or employment decisions.

Federal legislation begun during the Clinton administration and modified during the Bush administration has addressed some of these issues. The Health Insurance Portability and Accountability Act of 1996 (HIPAA, modified in 2002) requires a single signed consent to allow access in three domains: "treatment" (provision of health care and related services), "payment" (submitting insurance claims and determining eligibility), and "health care operations" (broadly defined sharing of information). The bill has been criticized for being coercive (people in health plans often must sign to receive treatment) and for the vague and broad scope of the third provision. Others have commended the plan for bringing a national uniform rule to these issues and for increasing access as well as control. For example, it is now easier for doctors to share patients' medical records, and there are fewer insurance approvals to sign (due to the single signature requirement). Control over information is more specific under this rule. Psychological and psychiatric records cannot be shared without a separate signature specifically for that information. With respect to test results, doctors must now ask how to contact the patient and whether leaving a message is acceptable. Also, without explicit permission, hospitals cannot even disclose whether someone is a patient, let alone disclose health status.

It has taken many years and much discussion to develop and enact HIPPA. Balancing access with confidentiality is a difficult task, especially when widespread use of computerized databases makes it easy to share confidential information and easy to inadvertently or deliberately share information that should be protected. Some patient advocates consider HIPPA to favor the medical system over individuals' rights to privacy, and reform efforts are under way. This is an issue that will receive ongoing attention as people seek

to achieve desired openness and closedness under changing circumstances.

5. SUMMARY AND CONCLUSIONS

5.1. Overview of Importance of Privacy Regulation

Privacy regulation is the selective control of access to the self or group. All societies have developed strategies through which their members can regulate interactions with others, although each society sets different standards of openness and closedness and has evolved different mechanisms for regulating privacy.

Privacy regulation involves many different kinds of events—from people standing too close in subways to burglaries being committed in neighborhoods, from concealing information about one's test scores to lying about medical conditions to obtain insurance. Privacy regulation also involves managing relationships to achieve a desired level of intimacy, perhaps by disclosing secrets to deepen a relationship or by avoiding listening to disclosures to maintain the relationship at a less intimate level. Privacy regulation involves choosing the levels of openness and closedness that one desires and then trying to achieve those levels by activating various privacy regulation mechanisms. Individuals and groups that can regulate their privacy are more viable and more likely to achieve their goals. When desired privacy is not achieved, people often complain of stress.

5.2. Prospects for the Future

Privacy regulation has been practiced and refined for centuries, yet people still have difficulty in managing interactions. This is due in part to differences between individuals in expectations about how open people should be and how to convey this through privacy regulation mechanisms. These problems occur in part due to cultural differences. As people immigrate to new countries, residents and new immigrants must recognize differences in privacy goals and strategies. Possibly, once people are aware of these differences, they can work on accommodating each other's preferences or on recognizing the cues used to signal whether the other wishes to be more open or more closed. In a changing situation, privacy regulation mechanisms and desires can also be changed.

Another source of problems is the rapidly changing technological environment and pressures to use technology for more openness. Psychological research on privacy suggests the following principles for guiding the use of these technologies:

1. *People should have reasonable control over others' access to them and people's access to others.* People should be able to proactively and reactively open and close themselves to others as they desire and see fit. This principle highlights privacy as a dialectic process of opening and closing rather than closing alone. It also emphasizes the importance of both a proactive and a reactive control capability. People should be able to proactively initiate openness or closedness rather than just react passively to the world.
2. *People may need a new or enhanced repertoire of regulatory mechanisms to control openness and closedness during the technological era.* Traditional face-to-face communication uses particular verbal, paraverbal, environmental, and cultural mechanisms. Although some of these mechanisms may apply during the technological arena, additional mechanisms—unique to new modes of communication—might be necessary.
3. *A dynamic privacy regulation system, with the ability to shift desired openness and closedness as circumstances change, will continue to be necessary and may require innovative applications of technology.* For example, some people like the anonymity of chat rooms, whereas others fear being lied to and deceived. To address this, some Web chat rooms may continue to be wide open and allow people to adopt personas and play out fantasies. In other chat rooms, participants may insist on knowing background information on each other and may agree to post authenticated vitae as the price for participating.
4. *Awareness of the importance of privacy regulation must be at the forefront of technological innovations.* Mechanisms that allow flexibility in achieving openness and closedness will be most successful.

See Also the Following Articles

Environmental Stress ■ Personal Space ■ Residential Satisfaction and Perceived Urban Quality

Further Reading

Altman, I. (1975). *The environment and social behavior: Privacy, personal space, territory, crowding.* Pacific Grove, CA: Brooks/Cole.

Altman, I., & Chemers, M. M. (1980). *Culture and environment.* Pacific Grove, CA: Brooks/Cole.

Altman, I., Vinsel, A., & Brown, B. B. (1981). Dialectic conceptions in social psychology: An application to social

penetration and privacy regulation. *Advances in Experimental Social Psychology*, *14*, 107–160.

Baum, A., & Epstein, Y. (Eds.) (1978). *Human response to crowding*. Hillsdale, NJ: Lawrence Erlbaum.

Baum, A., & Valins, S. (1979). Architectural mediation of residential density and control: Crowding and the regulation of social contact. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 12, pp. 131–175). New York: Academic Press.

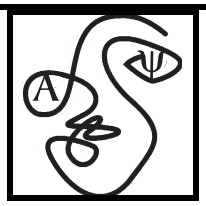
Brown, B. B. (1987). Territoriality. In D. Stokols, & I. Altman (Eds.), *Handbook of environmental psychology* (Vol. 1, pp. 505–531). New York: John Wiley.

Evans, G., Lepore, S. J., & Schroeder, A. (1996). The role of interior design elements in human responses to crowding. *Journal of Personality and Social Psychology*, *70*, 41–46.

Lepore, S. J., Evans, G. W., & Schneider, M. L. (1991). Dynamic role of social support in the link between chronic stress and psychological distress. *Journal of Personality and Social Psychology*, *61*, 899–909.

Margulis, S. T. (Ed.) (2003). Contemporary perspectives on privacy: Social, psychological, political [special issue]. *Journal of Social Issues*, *59*(2).

Petronio, S. (2002). *Boundaries of privacy: Dialectics of disclosure*. Albany: State University of New York Press.



Productivity

Robert D. Pritchard
University of Central Florida, Orlando, Florida, USA

1. Introduction
2. Defining Productivity
3. The Importance of Productivity
4. Improving Productivity
5. Approaches to Productivity Improvement
6. Measuring Productivity
7. Purpose of Measurement
8. Specific Approaches to Measuring and Improving Productivity
   Further Reading

GLOSSARY

effectiveness A measure of outputs compared with goals or standards; examples of effectiveness measures would include actual billable hours divided by the standard or expected amount of billable hours and actual percentage of returning customers compared with the goal of returning customers.

efficiency A measure of outputs divided by inputs; outputs are the products or services of the organization such as radios produced, customers satisfied, and/or clients seen, whereas inputs can be all inputs needed to produce the outputs such as labor, equipment, supplies, energy, and buildings.

feedback Information provided about employee behavior and outputs; it can be in many forms, including providing information on the amount or quality of work completed, subjective ratings of performance by managers, information provided informally by peers and subordinates, and a person's own perceptions about his or her work done.

goal setting A process whereby the subordinate and his or her supervisor jointly agree on a specific quantitative level of output that the subordinate will produce during the next work period (e.g., increasing revenues by 4% in the next quarter); less formal goal setting might be plans to make improvements or general intentions to perform better in the future.

incentives Outcomes of value that are provided to employees based on their levels of performance; most common are monetary outcomes (e.g., bonuses, incentive pay) and outcomes that have monetary value (e.g., vacations, sports equipment).

performance A general term that can refer to the amount or quality of the outputs produced, an evaluation of the outputs produced, or some combination thereof; it can be applied to an individual or to a much larger unit such as the total organization.

performance appraisal A process whereby one's supervisor, and sometimes one's peers and subordinates, evaluate the performance of an individual and provide this evaluation as feedback to that person; it is typically done with a series of ratings on various dimensions of performance such as technical knowledge and oral communications skills, but it can also include objective data on the quantity or quality of work done by the individual.

performance management A general term that includes all techniques used to influence and improve the performance of employees in organizations; such techniques include training, feedback, control systems, goal setting, and incentive systems.

productivity A term that can be defined in multiple ways, with the most traditional being outputs divided by inputs, usually expressed in monetary terms; it is also commonly defined as efficiency (outputs over inputs) plus effectiveness (outputs relative to a standard or goal).

©2004 Elsevier Inc.
All rights reserved.

productivity improvement Improvement in productivity measures over time.

role clarification Techniques used to make the expectations of the organization clear to the individual; examples include discussions about expectations with one's supervisor, formal systems that identify the relative importance of different parts of the work, and information systems that indicate what the priorities are for making performance improvements.

Productivity is an important topic, and the field of psychology has made many contributions to understanding and improving productivity in organizations. Productivity, especially improving productivity, has important benefits ranging from global effects to effects on individuals' quality of life. There are a number of different definitions of productivity, and it is important for communicating with others to be clear on what definition is being used. Psychologists focus on improving productivity by influencing people's behavior, especially through techniques such as feedback, motivation, goal setting, incentives, and role clarity. Measuring productivity is complex, and this article describes both the purpose of the measurement and the criteria for good measurement. Once measured, productivity is normally used as feedback, so the article summarizes criteria for optimal feedback design. Finally, the article notes specific approaches to measuring and improving productivity.

1. INTRODUCTION

1.1. History

Labor economists originally used productivity as a formal term during the early 1900s. During the 1950s and 1960s, the United States was the most productive country in the world, and there was little doubt that this state of affairs would continue. However, by the 1970s, American superiority in productivity was being threatened by productivity in other countries, and this was seen as a major concern that would lead to serious long-range economic problems. For example, the success of the Japanese in U.S. automobile and steel markets led to plant closings and lost jobs, major U.S. companies were in severe financial difficulties, and what had historically been a trade surplus became a trade deficit that was increasing.

1.2. Productivity as a National Issue

Although many causes were responsible for these problems, loss of clear U.S. superiority in productivity was seen as one of the major ones. Because of this, productivity and productivity improvement became issues of great social and political importance, not only in the United States but in all industrialized nations. This led many other disciplines, besides economics, to focus on productivity issues, and one such discipline was psychology.

2. DEFINING PRODUCTIVITY

Defining productivity is a difficult task. Since the 1970s, behavioral scientists of many types have worked in the area of productivity, resulting in a proliferation of definitions of productivity. However, the vast majority of definitions fall into one of three categories.

2.1. Efficiency

The first category defines productivity as an efficiency measure where outputs are divided by inputs. Thus, the number of outputs generated, such as refrigerators, potatoes, or university graduates, is divided by the inputs used. Such inputs usually include raw materials, labor, energy, and so forth. Typically, both outputs and inputs are measured in terms of cost, so that the two can be divided to form a ratio that is interpretable across different settings. This definition is used by most economists and by some behavioral scientists.

2.2. Combinations of Efficiency and Effectiveness

The second definition of productivity is a combination of efficiency and effectiveness. Efficiency is the output/input measure used in the first definition. Effectiveness is the ratio of outputs relative to some standard or objective. For example, the number of trucks produced relative to a specific quantitative goal is an effectiveness measure. Those taking the approach that productivity should include both efficiency and effectiveness argue that both are necessary for a complete definition of productivity. One such definition is that productivity refers to how well an organization uses its resources to achieve its objectives.

2.3. Broader Definitions

The third type of definition is the most broad one and essentially includes as productivity any characteristic that makes the organization function better. This approach includes efficiency and effectiveness but also includes factors such as quality of output, work disruptions, absenteeism, turnover, and customer satisfaction.

2.4. Definitions and Clear Communications

It is important for anyone working in the area of productivity to be aware that these different definitions exist. Some researchers and practitioners are very rigid in their choice of definition and do not accept that the other definitions are "productivity." Thus, for effective communication, the same definition of productivity should be used by all.

3. THE IMPORTANCE OF PRODUCTIVITY

Productivity and productivity improvement are important because they influence many aspects of our lives. They have effects that range from global effects to those on specific individuals.

3.1. Global and National Benefits

Increased productivity means generating the same goods and services with less inputs. Thus, it is a way in which to conserve societal or global resources ranging from trees to human labor. Increased productivity also allows for more outputs to be available for the same inputs. Thus, we can be closer to a global society of plenty while using less of our global resources. At the national level, there are clear ties between productivity and important economic outcomes such as controlling inflation. For example, the real cost of goods is influenced by productivity because productivity improvement results in producing the same goods for lower costs. Assuming a roughly constant profit margin, the real cost of goods decreases as productivity increases.

3.2. Benefits for Industries and Specific Organizations

Productivity and productivity improvement are also important at the levels of the industry and the individual firm. If productivity improvement of an industry or a firm is higher than that of its competitors, that industry or firm will more likely survive and prosper. Costs and prices are lower, thereby making those products and services more competitive in the marketplace. This leads to higher sales, greater profits, and more jobs.

3.3. Benefits for Individuals

Productivity improvement can also increase the quality of our lives. Productivity gains will lead to efficient use of our time and more leisure time, and high productivity is a key to advancement in organizations. Equally important, people like to be productive. It is a central aspect of self-fulfillment and self-respect.

4. IMPROVING PRODUCTIVITY

The major reason why people focus on productivity is to improve productivity. It is by improving productivity that the gains from the global level to the individual level can occur.

5. APPROACHES TO PRODUCTIVITY IMPROVEMENT

There have been two major approaches to improving productivity. The first is through a change in technology. Some economists argue that most or all major improvements in productivity can be traced to technological improvements. The other approach is behavioral. Behavioral scientists argue that new technology can improve productivity only if organizational personnel use it effectively; thus, technological improvements also require a behavioral component. Furthermore, behavioral scientists argue for the importance of issues such as performance measurement, motivation, feedback, goal setting, incentives, and role clarity. It is here that the field of psychology has made the greatest contributions.

6. MEASURING PRODUCTIVITY

To improve productivity, it is important not only to measure it but to measure it well. However, productivity measurement is very complex, and a number of issues must be considered in designing a productivity measurement system.

7. PURPOSE OF MEASUREMENT

The first step is to decide on the purpose of the measurement because that decision will affect how the measurement is done. One purpose in measuring productivity is to compare large aggregations of organizations with each other. Examples of this purpose include comparing the United Kingdom's economy with Japan's economy and comparing the electronics industry with the automobile industry. Another purpose is to evaluate the overall productivity of individual organizations for comparison with each other or with some standard. A third purpose is as a management information system or control system. Here the focus is on a single organization, and the measurement deals with the functioning of the human/technological system. Finally, productivity could be measured for use as a motivational tool. This is the typical purpose of the organizational psychologist. The assumption is that if individuals change their behavior appropriately, productivity will increase.

7.1. Criteria for a Good Productivity Measurement System

There are many specific issues that must be considered in measuring productivity by the psychologist or organizational behavior specialist. These include ensuring that all important aspects of the work are measured, having reliable and valid measures that are cost-effective to collect, using measures that are understood by all, using measures over which the personnel have control, ensuring that the measurement system will be accepted by both the managers and the people doing the actual work, and ensuring that measures for each unit are aligned with the strategic objectives of the broader organization.

In most approaches to productivity improvement, the productivity measurement system is used for feedback to the personnel doing the work. The assumption is that such feedback will help to produce the behaviors that will increase productivity. However, to be most effective, the feedback system must meet an additional series of criteria beyond those for the productivity measurement system. Feedback must be given in a timely manner, be understood by those receiving the feedback, contain both descriptive information (e.g., the unit finished 15 orders) and evaluative information (e.g., 15 orders is very high performance), give information on all important aspects of the work, account for the varying importance of different aspects of the work, provide an overall productivity score, identify improvement priorities, and allow for unit personnel to plan new strategies for doing the work.

Research has indicated that such feedback can be very effective in improving productivity. For example, one approach developed by Pritchard and associates measures productivity and uses these measures for feedback. It is called the Productivity Measurement and Enhancement System or ProMES. Figure 1 shows the effects of feedback due to ProMES as shown from field studies on 55 organizational units located in seven different countries. The nature of the organizations varied, and the research was conducted by nine different research groups. The horizontal axis represents time beginning with baseline (prior to the use of feedback) and continuing through the implementation of feedback. The vertical axis represents the overall effectiveness score, which is an index of overall productivity. Once feedback was introduced, there were large increases in productivity.

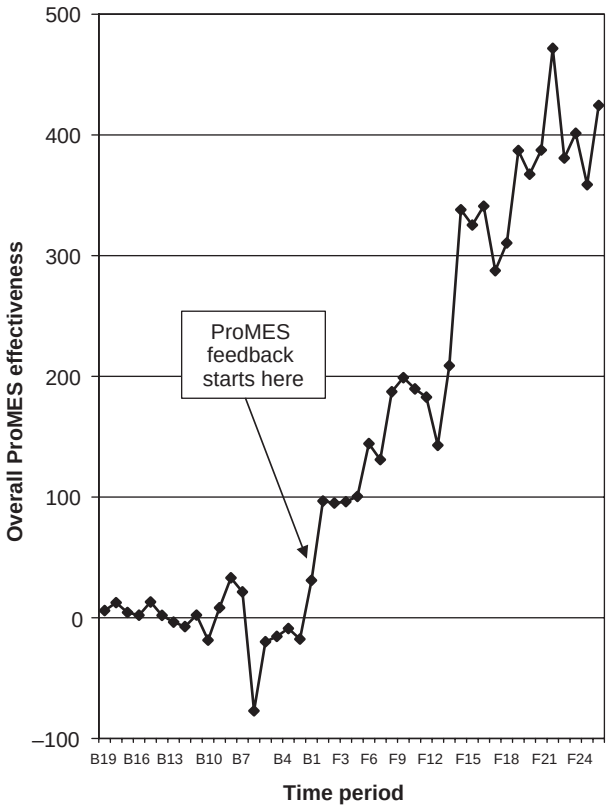


FIGURE 1 Effects of ProMES feedback on productivity. Reprinted from Pritchard and colleagues (2002).

8. SPECIFIC APPROACHES TO MEASURING AND IMPROVING PRODUCTIVITY

Psychologists and other organizational specialists have proposed a number of other specific techniques for measuring productivity that focus on using productivity measures as feedback to help personnel improve their productivity. The best-known examples are the Balanced Scorecard and Six Sigma.

The Balanced Scorecard was originally designed as an aid to top management. It stressed that other measures besides financial measures are important to overall organizational effectiveness and that including these nonfinancial measures would result in a more "balanced" evaluation. It has since been broadened and is now used for all levels of the organization. Six Sigma was originally developed for feedback and productivity improvement in manufacturing settings. The central idea is that processes with multiple steps, such as assembling cars and refining petroleum, require that each step be done with extremely high reliability for the final product to meet required specifications. Measures of each step in the process are taken, and efforts are made to increase the reliability of each step.

Finally, there is a class of practices that are designed to improve performance and thereby productivity. These practices, which have come to be called "performance management," include performance appraisal, other forms of feedback, goal setting, incentives, and role clarification.

See Also the Following Articles

Competence at Work ■ Groups, Productivity within ■ Work Motivation

Further Reading

Campbell, J. P., & Campbell, R. J. (Eds.). (1988). *Productivity in organizations*. San Francisco: Jossey–Bass.

Dasgupta, T. (2003). Using the Six Sigma metric to measure and improve the performance of a supply chain. *Total Quality Management*, *14*, 355–366.

Guzzo, R. A., Jette, R. D., & Katzell, R. A. (1985). The effects of psychologically based intervention programs on worker productivity: A meta-analysis. *Personnel Psychology*, *38*, 275–291.

Kaplan, R. S., & Norton, D. P. (1996). *Translating strategy into action: The Balanced Scorecard*. Boston: Harvard Business School Press.

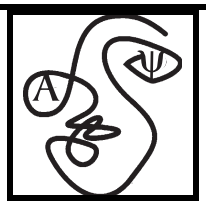
Kendrick, J. W. (1984). *Improving company productivity*. Baltimore, MD: Johns Hopkins University Press.

Locke, E. A., & Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American Psychologist*, *57*, 705–717.

Pritchard, R. D. (1992). Organizational productivity. In M. D. Dunnette, & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (2nd ed., Vol. 3, pp. 443–471). Palo Alto, CA: Consulting Psychologists Press.

Pritchard, R. D., Holling, H., Lammers, F., & Clark, B. D. (Eds.). (2002). *Improving organizational performance with the Productivity Measurement and Enhancement System: An international collaboration*. Huntington, NY: Nova Science.

Sink, D. S. (1985). *Productivity management: Planning, measurement and evaluation, control and improvement*. New York: John Wiley.



Pro-environmental Attitudes and Behavioral Change

Henk Staats

Leiden University, Leiden, The Netherlands

1. Introduction
2. Environmental Problems as Social Dilemmas
3. The Relation between Attitudes and Behavior
4. How to Effect Pro-environmental Behavior Change: The Role of Psychology
5. Considerations Regarding the Effectiveness of Interventions for Pro-environmental Behavior Change
6. Conclusions
Further Reading

GLOSSARY

attitude The general evaluation of an object; in the theories described in this article, it refers to the general evaluation of a behavior.

intervention technique A method to change behavior in a certain direction; in this article, it refers to a method to change antienvironmental behavior to pro-environmental behavior.

pro-environmental behavior Behavior that is relatively favorable for the environment in comparison with behavior that serves the same primary function (e.g., using public transportation instead of driving a car).

social dilemma The tension between individual interests and collective interests.

To understand pro-environmental behavior and be able to influence it, it is essential to apply knowledge from attitude–behavior theories. However, the central characteristic of pro-environmental behavior can be understood only if knowledge from social dilemma theory is applied simultaneously.

1. INTRODUCTION

Large-scale environmental problems have been on the political agenda for more than 30 years but continue to threaten the earth's inhabitants. Clearly, some of the problems are particularly stubborn and defy easy solutions. Basic to an understanding of environmental problems is the fact that they are ultimately phenomena created by human behavior and, thus, are to be solved by human behavior. Gradually, this has become recognized in a field that was traditionally dominated by experts from the physical sciences such as chemists, biologists, and ecologists.

This article gives an overview of contributions from social/environmental psychology that deepen insights into the human dimension of environmental problems and that are helpful in designing interventions that change behavior in a pro-environmental direction. First, the article describes a number of research traditions that are highly relevant for understanding behavior that affects the environment. These deal with attitude–behavior relations, moral considerations, and

©2004 Elsevier Inc.
All rights reserved.

habits. Overarching these perspectives is the conceptual framework of social dilemmas, a paradigm that is essential for understanding the nature of environmental problems. The article then describes what psychology has to contribute when considering ways of changing behavior in a pro-environmental direction.

2. ENVIRONMENTAL PROBLEMS AS SOCIAL DILEMMAS

Essential in thinking about the origin of many environmental problems is their social nature, originating from a fundamental tension between acting in the service of direct individual interests and acting in the service of long-term community interests. Many societal issues have this structure, and it is almost essential to an organized collective that individuals sacrifice, to some degree, their individual interests in favor of the interests of the community. Consider, for example, the principle of paying tax to finance collective institutions and services such as police, health care, and infrastructure for transportation. From an individual perspective, it would be profitable to evade the obligation of paying taxes. But when everyone would succeed in doing this, no public good would become available and everybody would be worse off. This is the core of the social dilemma that lies at the heart of many environmental problems. It was described in 1968 by Hardin as "The Tragedy of the Commons." That publication has become emblematic for the analysis of environmental problems.

Different in comparison with the taxpayers' dilemma is that environmental problems are often delayed in time and that effects are not necessarily manifest at the place where the problem is created. Consider the greenhouse effect, one of the most serious environmental problems of our time. One reason why the greenhouse effect is difficult to solve is that effects are expected to be delayed some 50 years. Also, it will become most threatening in parts of the world that have hardly contributed to the problem but that will be most vulnerable to the expected rise in sea level.

The social dilemma paradigm has proven to be a fruitful area of social–psychological experimental research, providing information about which factors are likely to influence the trade-off between choosing for the private interest and choosing for the collective. Important factors include the payoff structure (i.e., how much the individual loses by acting in the collective's interest), communication about the decision among group members, expectations about the others' cooperation, perceived efficacy (i.e., how important the individual contribution is in determining the outcome of the group), and identifiability (i.e., the degree to which a person is known to have decided in favor of individual interests vs collective interests).

3. THE RELATION BETWEEN ATTITUDES AND BEHAVIOR

The social dilemma framework helps one to understand how behavior can be differentially driven toward individually valued outcomes or collectively valued outcomes such as the quality of the environment. However, it gives only a limited account of how behavior results from cognitive and motivational processes. Three theories are specifically important in the context of environmental behavior.

3.1. Theory of Reasoned Action

One of the most influential models in the social–psychological literature is the theory of reasoned action (TRA) developed by Fishbein and Ajzen. The TRA postulates that behavior is the result of three main factors: attitude toward behavior, subjective norm, and behavioral intention. Attitude and subjective norm are postulated to cause behavioral intention, which in turn causes behavior.

In the TRA, Fishbein and Ajzen confined the attitude concept to the general evaluation of a behavior. Other theories have no restrictions as to the nature of the attitude object and, thus, may consider ideas, persons, things, issues, ideas, or other entities. However, these more general theories have fared less well when it was important to understand and predict behavior. Therefore, at the time, Fishbein and Ajzen's theory came as a much-needed innovation in the domain of attitude theories.

The attitude is measured with a number of adjectives (e.g., good–bad, pleasant–unpleasant, positive–negative) that are generally statistically combined to obtain a general judgment of the performance of a certain act.

Subjective norm is the perception of the individual that, in general, other people who are important to the individual want him or her to perform, or not to perform, that behavior. Who is important may vary across behaviors, but often this will include family, friends,

and neighbors. Other influences may stem from a religious institution to which a person feels attached, a political party, and the environmental movement. It is the individual's overall impression of how to behave, according to this collective of reference persons and groups, that is defined as the subjective norm.

The third concept is behavioral intention, that is, the deliberate plan to perform the behavior. According to the theory, the intention is the result of a decision process based on the attitude and the subjective norm, and it is the best predictor of behavior.

Finally, behavior is the performance of a certain act, In the TRA, behavior is usually confined to acts that are observable by others such as separating garbage and buying ecological food.

3.1.1. The Backgrounds of Attitude and Subjective Norm

In the TRA, the concepts of attitude and subjective norm both are explained by a combination of two elementary concepts. The attitude toward a behavior is caused by beliefs about the likelihood of salient outcomes of performing a behavior, weighted by the evaluation of each of those outcomes. Salient outcomes are the outcomes that immediately come to the mind of the individual when thinking of the behavior. This combination of belief strength and evaluation of each outcome was formalized by Fishbein and Ajzen in the following formula: $A_{act} = \Sigma B \times E$. The formula states that the attitude toward a certain behavior (A_{act}) is the sum of the products of beliefs (subjective estimate of likelihood of occurrence) and evaluations (valence) of all the behavioral outcomes that are considered by the individual.

Subjective norm, like the attitude concept, is composed of two subconcepts that explain which persons and groups are responsible for the normative pressure to perform, or not to perform, a certain behavior. The first concept is that of normative belief. A normative belief indicates the extent to which a person thinks that a specific referent person or group wants him or her to perform a behavior. Usually, with each behavior, the individual holds a number of normative beliefs. Referent persons will be family members (e.g., spouse, children, parents) who are important for many behaviors, whereas the influence of other persons or groups may vary more strongly with the behavior. For some behaviors, such as recycling of organic waste, the opinion of an environmental organization may be quite influential, whereas the opinion attributed to this group will be irrelevant when deciding whether to make a visit to one's parents' house over the weekend. Each normative belief (*NB*) is weighted for the so-called motivation to comply (*MC*), that is, the degree to which an individual allows this referent person or group to exert influence on him or her. Subjective norm (*SN*) is the sum of the products of all normative beliefs and the corresponding motivation to comply. Expressed as a formula, $SN = \Sigma NB \times MC$.

3.1.2. Correspondence: A Major Condition for Application of the TRA

Correspondence entails that all of the components of the model are measured at the same level of specificity. Correspondence must exist on four criteria: action, target, context, and time. Action refers to the behavior itself. For example, if the behavior that is being investigated is recycling, all concepts of the theory must be specified as recycling and not as recycling in the measure of behavioral intention and as pro-environmental behavior in the measurement of the attitude. Similarly, one should specify the target (also called object sometimes) consistently. If one is interested in predicting the recycling of organic waste and not of used paper, aluminum cans, or glass, one should specify each component of the model as pertaining to the recycling of organic waste. And again, concerning context, one should decide whether one wants to investigate people's recycling behavior of organic waste in their homes, in their workplaces, or perhaps in all of the places where they spend time. Finally, concerning time, one must decide whether one wants to investigate people's recycling of organic waste next week (e.g., because there will be a special organic waste recycling drive organized by an environmental organization), the coming year, or forever from now on. Note that it is a matter of choice by the investigator at which level of specificity all of the components of the model are measured. Note also that the often used concept of environmental concern can be considered to be a very general attitude that is not correspondent with specific behavioral measures. Therefore, relationships between environmental concern and specific behaviors will necessarily be weak.

3.2. The Theory of Planned Behavior

The TRA assumes that the behavior under investigation is under volitional control, that is, that people believe that they can execute the behavior whenever they are

willing to do so. Gradually, the TRA was used more often for the study of behaviors for which control was a variable factor. For that purpose, the TRA was complemented by a component that was named perceived behavioral control. This concept represents the extent to which people believe they are able to perform the behavior because they have adequate capabilities and/or opportunities or are lacking in these. It is very easy to see that this factor can substantially improve the generality of application of the model because there are many behaviors that need specific skills or external facilities. For example, recycling is virtually impossible if no collection system is available, and abandoning private cars is often impractical, at the least, when public transportation functions poorly.

The successor of the TRA that incorporates volitional control is the theory of planned behavior (TPB). Similar to attitude and subjective norm, the joint effects of two subconcepts determine behavioral control. The first subconcept consists of so-called control beliefs, that is, the estimated likelihood that each of a number of specific factors will facilitate or impede execution of the behavior. An example might be the following control belief: "I can go to my office by bus given the distance from the bus stop to my home." The second factor is called perceived power, that is, a judgment of the degree of facilitation or impediment that each specific control belief represents. An example might be the following: "The distance from the bus stop to my home makes it very easy/very difficult to go to my office by bus." The formula combining the two subconcepts of likelihood of control and perceived power is identical to that of the attitude and subjective norm concepts: $\mathrm{PBC} = \Sigma C \times P$.

Both the TRA and the TPB have been used repeatedly for investigating specific environmental behaviors such as changing travel mode, water conservation, recycling, and green consumerism. In general, the models have proven to be useful in understanding the behavior, with important contributions of perceived behavioral control.

3.3. Norm Activation Theory

The attitude concept, as used in the TPB, is meant to be a general evaluation of a behavior. This implies that all of the specific evaluative criteria that guide a decision to adopt, or not to adopt, that behavior should determine the attitude score. This does not seem to be the case given that research has shown repeatedly that moral considerations are not effectively covered by the attitude concept. Environmental behavior, because of its social dilemma character, clearly has a moral dimension in that acting in favor of the environment/community is generally perceived as morally superior to acting out of self-interest. The development of moral considerations and their relation to behavior is described in Schwartz's norm activation theory (NAT). The central concept is personal norm, defined as intrinsically motivated self-expectations regarding morally appropriate behaviors. Norm activation generates a feeling of personal obligation to perform a specific behavior. According to the NAT, norm activation occurs under the influence of four situational activators and two personality trait activators. Norm activation starts with the awareness of need. It involves the extent to which a person's attention is focused on the existence of a person or a more abstract entity (e.g., the environment) in need. Apart from being aware of need, the potential actor must define the situation at hand as one in which he or she feels some responsibility for the consequences of the needy party's welfare. Norm activation is further enhanced by efficacy, that is, the extent to which a person recognizes actions that might alleviate the need. The actions that are recognized as such determine which specific norms are activated. Finally, ability refers to the actor's perception about whether he or she possesses the resources or capabilities needed to perform the focal behavior.

In addition, the NAT proposes that two personality traits influence activation of personal norm: awareness of consequences and denial of responsibility. Awareness of consequences refers to a person's receptivity to situational cues of need. Denial of responsibility refers to a person's inclination to deny responsibility for the consequences of his or her behavioral choices for the welfare of others. The four situational activators and the two personality traits determine whether or not a behaviorally specific personal norm becomes activated. Specific behaviors, such as burning garden waste, recycling, buying lead-free gasoline, using off-road vehicles in the environment, energy conservation, and littering, have been studied by means of the NAT. In general, just parts of the whole model have been used. Most studies show that personal norms significantly influence pro-environmental behavior, whereas there is less agreement about the predictors of personal norms and the exact process through which personal norms become activated.

3.4. Habits

All of the approaches to understanding environmental behavior described in this section assume that people process relevant information to decide to perform an environmentally relevant behavior in a pro- or antienvironmental manner. However, people often rely on behavioral routines or habits. Habits are especially likely to develop when behaviors are displayed frequently and in a stable context. Habitual behavior is defined as behavior that is displayed automatically on the presence of a goal, that is, a direct goal–action link that is not preceded by consciously developed intentions. Automatically executed behavior has three main characteristics in that (a) it can be executed fast, (b) its execution cannot be controlled, and (c) it is very efficient (i.e., it allows other cognitive processes to be executed simultaneously). Care should be taken not to equate habitually performed behavior with all behaviors that are performed in the past, even when the conditions of high frequency and context stability are met. Behavior executed in the past can influence future behavior through its effect on intentions, and in that case it is unlikely that the behavior is a habit. However, some studies have shown that the strength of intentions to influence behavior depends on the degree to which behavior has been performed in the past. When behavior has been frequently performed in the past, the effect of intentions on behavior is smaller. In such cases, where past behavior changes the intention–behavior relationship, it is very likely that past behavior indicates that habits play a role.

Clearly, the automatic execution of behavior has important advantages. But it has drawbacks as well in that people who have established habits pay less attention to information that might be important for changing behavior. For example, the information that is issued by a railway company about improved services is likely to be ignored by people with strong car travel habits, even though it might be very advantageous for them. Given that people with strong habits cannot easily be influenced by information, an important question then becomes how to change habitual antienvironmental behavior.

4. HOW TO EFFECT PRO-ENVIRONMENTAL BEHAVIOR CHANGE: THE ROLE OF PSYCHOLOGY

Applied psychology can play a role in establishing more environmentally sound living patterns. This section gives a summary of the potential contribution of psychology for the development of intervention methods. The methods range from obligatory to voluntary systems, but all need the application of psychological knowledge to understand and optimize effects.

4.1. Laws, Rules, and Regulations

It seems to be a straightforward solution to communal problems to implement rules, regulations, or laws that prohibit the actions that cause the problem. However, from a psychological perspective, there are several issues that need mentioning. First, regulations will fail when the targeted population is not convinced of the necessity of the policy measure. In that case, strict enforcement will be necessary and costly. If strict enforcement is not applied, the targeted population will not adhere to the regulation. In both cases, the measure will not have the favorable effects that are desired. Second, a policy measure will be accepted according to its perceived fairness. Fairness means that people will realize that the policy measure targets the right actions and punishes offenders according to the severity of their offenses. Third, from these characteristics and based on laboratory research on social dilemmas, it can be predicted that people will be most willing to endorse policy measures when they are in favor of the proposed measure but are very uncertain what other people's attitudes and behavior will be. These psychological effects of perceptions, attitudes, and uncertainties and their relationships were formalized in the structural goal expectation theory developed by Yamagishi in 1986. Fourth, implementation of policy measures, even if they are successful, is not an easy way out of the social dilemma. To develop legislation, political support for measures protecting the environment is necessary. To create political support, environmental awareness and concern among voters, and willingness to sacrifice personal advantages for long-term collective benefits are necessary.

4.2. Changes in the Physical Environment

Measures that are generally less dependent on political approval are changes in the physical environment that facilitate pro-environmental behavior or impede antienvironmental behavior. These kinds of measures can work in three different ways.

4.2.1. Direct Influence on Behavior

Changing the layout of a road (e.g., by creating additional curves or narrowing lanes) will reduce automobile speed and emissions directly.

4.2.2. Influence on Habits That Control Behavior

Many environmental behaviors are habitually performed. Changing the physical context of the behavior means that these habits no longer apply and information must be sought about how to achieve the behavior's goal in these changed circumstances. This renewed need for information creates possibilities for pro-environmental behavior change, for example, by showing that the daily commute to work can now be completed faster by public transportation than was the case during the period when the car commute habit was developed.

4.2.3. Influence on the Convenience of Performing Pro-environmental Behavior

In many cases, pro-environmental behavior is impeded not by negative attitudes but rather by weak positive attitudes combined with a lack of perceived behavioral control. For example, most people are in favor of recycling garbage but are discouraged by the time and effort it takes to bring their recyclable garbage to a recycling station. Creating drop-off facilities for recycling paper, bottles, and other recyclables at close range from home may strongly increase the volume of recyclable materials.

4.3. Monetary Incentives

It might seem as though the most straightforward way in which to solve environmental problems is to let polluters pay for the damage they do to the environment. In general, this may be an effective strategy that eliminates the social dilemma structure of environmental problems. Research has shown that people are indeed sensitive to changes in the payoffs and that the number of egoistic choices (as compared with choices for the collectivity) is reduced when egoistic choices are less profitable. However, the following subsections review a number of complicating factors pertaining to the psychology of justice, the perception of monetary changes, and the dependence of monetary incentives on information to accomplish the desired effect. Moreover, financial measures run the risk of weakening intrinsic motivation, that is, the motivation to act in a pro-environmental manner out of deeply felt concern for the state of the environment.

4.3.1. Policy Measures

Financial measures, in particular, can be evaluated differently depending on the fairness principle that is considered appropriate. Three principles are applied: those of equity, equality, and need. Equity implies that the damage is compensated by the actor in proportion to the damage that he or she causes. Equality implies that people contribute to the same extent, for example, that they pay the same amount for having their houses connected to the city's sewer system regardless of the amount of polluted water they produce. Need implies that people with more financial resources assume a larger share of the financial burden. It is not always clear which principle should apply in a specific situation. This political disagreement constitutes a reason for the absence or delay of appropriate measures.

4.3.2. Price Perception

Price perception is subject to many different influences that can substantially change the relation between actual costs and willingness to take pro-environmental action. A finding that is at the core of Tversky and Kahneman's prospect theory is that people are more sensitive to potential losses than to potential gains. For that reason, it is advised to frame financial consequences of antienvironmental behavior as a potential loss instead of framing financial consequences of pro-environmental behavior as a potential gain. The complexity of consequences of financial measures may cause the use of heuristics. For example, the calculations necessary to decide whether it is cost-efficient to install double-glazing are difficult because it is necessary to simultaneously consider changing consumption, changing energy prices, costs of a loan, and/or lack of profit from a different investment. Consumers, and even entrepreneurs, are rarely able to take all of the relevant factors into account and may simply evaluate a decision on the basis of nominal costs before and after installment.

4.3.3. Financial Measures

Measures such as subsidies, loans, discounts, and other price changes do not automatically lead to changes in

behavior. Being aware of those price changes is crucial, and this implies the importance of informational interventions that precede or accompany the introduction of financial measures. Several experiments have demonstrated that attractive financial propositions are successfully implemented only when sufficient attention is given to making the programs known. To increase participation, it is often more prudent to spend money on information campaigns than to further increase the financial attractiveness of propositions.

4.3.4. Target Recurring Behavior

When preparing financial measures that target recurring behavior, one should bear in mind that the motivation to perform pro-environmental behavior is changed from intrinsic (out of concern for the environment) to extrinsic (for the monetary advantage). For one-time decisions, this might not be an issue. However, for recurring behavior, a relapse to antienvironmental behavior might occur once the financial gain from acting in a pro-environmental manner is removed. Pro-environmental behavioral maintenance then becomes problematic. Another unwanted effect that may stem from financially favorable energy-saving devices is the rebound effect. The knowledge that energy, or water, is so sparsely used due to new efficient devices might invite people to be less careful in their use of these devices. Having lights on throughout the night and showering longer might be invited by the installment of energy-saving lamps and water-saving showerheads, respectively. This phenomenon helps to explain why reductions in energy use through new efficient devices are generally not as favorable as foreseen in the calculations of technicians.

4.4. Information Strategies

As has now become clear, information is indispensable for implementing other interventions. It is helpful in increasing acceptance for policy measures, in announcing that physical changes in the environment are forthcoming, in announcing that financial support can be received for the implementation of energy-saving measures, and so on. Apart from supporting other measures, information can be used on its own for a large number of purposes, all ultimately intended to change behavior in a pro-environmental direction. Prominent functions of information include the following:

1. Information may create insight into the structure of the dilemma of environmental behavior. Laboratory research on social dilemmas has pointed out that people are more willing to cooperate when they are made aware of the dilemma structure of environmental problems.
2. Information may increase problem awareness and emotional concern. Both are important, but they are important to different degrees for different target groups. Groups of people who are motivated and capable of processing information thoroughly are better served by more factual information. Knowledge about the environmental problem—its nature, causes, and consequences—is an important condition to induce pro-environmental behavior, according to Schwarz's norm activation theory. For people who are relatively lacking in ability and motivation to process information, messages that stress the affective consequences of environmental degradation may be more influential.
3. Information can teach people relevant skills that facilitate the performance of pro-environmental behavior. People should learn how to dispose of chemical waste and how to compost organic waste before such pro-environmental behavior is applied. Many studies point out that the relative ease of performance is a crucial factor in pro-environmental behavior.
4. Information may tackle antienvironmental habits by reminding people of their pro-environmental attitudes. Using prompts to remind people not to drive fast, to lower thermostat settings well before going to bed, and so on can be successful in that these prompts encourage behaviors with which people agree.
5. Information may be used to address possible personal benefits of pro-environmental behavior. Sometimes, wrong estimates are made regarding the costs and benefits of pro-environmental behavior. This may occur, for example, when someone estimates the financial consequences of home insulation. Wrong estimates can be caused by outdated information (among other reasons), and this is especially likely in the case of habitual behavior when less attention is given to new information. This neglect of information may, for example, impede a choice for public transport even though travel conditions have improved greatly compared with the past when the choice to drive private automobiles was made.
6. Communication can reduce social uncertainty. Knowing that others are also willing to incur the costs of pro-environmental behavior convinces people that their efforts will not be in vain. This confidence in the efficacy of shared efforts will increase people's

willingness to behave in a pro-environmental manner. It should be noted, of course, that this mechanism works only when others (are willing to) behave pro-environmentally. If this is not the case, communication strategies should focus on other determinants such as knowledge, concern, and ability.

5. CONSIDERATIONS REGARDING THE EFFECTIVENESS OF INTERVENTIONS FOR PRO-ENVIRONMENTAL BEHAVIOR CHANGE

Changing behavior is sometimes a difficult task, and this certainly applies to changing environmental behavior in a pro-environmental direction. The social dilemma nature of environmental problems makes personal benefits of change, and thus of pro-environmental change, less likely. Nevertheless, people are willing to change to some extent. The categories of instruments described in the previous section may be effective in inducing change. But what does effectiveness actually mean? Contrary to what one would think intuitively, effectiveness is not a unitary concept. In what follows, this section adheres to the evaluation system proposed by De Young in 1993. According to De Young, effectiveness must be broken down into five dimensions:

- *Reliability*. Is an intervention able to effect the desired change in the target group? Moreover, is it able to do this every time? This criterion comes closest to what one would ask of an intervention program. However, an answer may be subtler than the questions imply. For example, the technique may be able to change individual behavior to some extent and may be more successful with some individuals than with others. The question about repeated use implies that the technique is not expected to create a lasting change that is exhibited each time the behavior can be displayed.
- *Speed of change*. How fast are changes in behavior accomplished? Of course, the sooner, the more impact (aggregated over time). However, some interventions can be expected to have their effect only after some time. The installation of expensive energy-saving equipment, such as a new central heating furnace, will usually be done when the old equipment is worn out. For research on effectiveness of intervention programs, it constitutes a challenge to be able to discover changes initiated by the program but effected over a longer term. However, this would be desirable to know so as to be able to give a good estimate of the value of an intervention program.
- *Particularism*. This is the degree to which an intervention must be tailored to the characteristics of a target group. This ranges from an intervention that is universally applicable to an intervention that is uniquely designed for an individual. Money is an example of a motivator that is considered to be generally applicable because it retains its value regardless of the person intervening, whereas personal attention seems to be at the other extreme because it clearly matters who is giving personal attention. An intervention that is universally applicable and effective is, of course, a great tool because it can be administered much more efficiently. However, in general, a degree of tailoring to characteristics of the target group is necessary to increase the effect of an intervention. Favorable in that respect is a recent development that allows messages to be tailored to the target person or group in an automatized process by which personalized information is sent to individuals whose characteristics have been inventoried earlier by means of a questionnaire and stored in an electronic database. Based on this information, a computer program automatically compiles an informative message that is closely tuned to the receiver. This technique combines the reach of mass media with the advantages of personalized advice; therefore, it seems promising.
- *Generality*. To what extent does an intervention technique influence behaviors that are not directly targeted but are conceptually related? And how closely related should they be to be influenced by the intervention? In general, it has proven to be difficult to create a degree of transfer from one behavior to the next. This is considered one of the two main challenges for an environmental psychology that aims to contribute to substantial changes in environmental behavior.
- *Durability*. To what degree are effects of intervention techniques lasting? Achieving durability of pro-environmental change constitutes the second main challenge. Especially with intervention techniques that focus on voluntary participation, it has been difficult to create substantial effects that last beyond the period when the intervention is executed. Examples of long-lasting changes are scarce. However, it seems that some techniques are more promising in this respect than are others. Especially explicit commitment (e.g., by a pledge done in public to

adhere to the goals of the intervention) seems to offer possibilities of behavioral maintenance after the intervention has expired.

6. CONCLUSIONS

Environmental problems are essentially problems of human behavior. Psychology can help one to understand behavior and propose ways of changing behavior. These are the main tenets of environmental psychology. This article has given a brief overview of influential theoretical perspectives and of the way in which psychology can help to create or improve pro-environmental behavior. It seems that a considerable knowledge base has been built over the past 30 years but that much more scientific attention, creativity, and persistence are needed, especially in developing interventions that change behavior. The rate of pro-environmental changes, and especially the durability of changes, is problematic. This is very understandable when considered from the perspective of social dilemmas. Given the existing structures in society, it is difficult to forgo individual interests for environmental improvements that are uncertain, may take a long time to come about, and may even be mainly beneficial to others. Possibly, psychology must dedicate more attention to the key actors in society—politicians, managers, and people who lead and represent large organizations. To this point, most attention has been on understanding and influencing people as consumers and citizens. For the most part, interventions have aimed at influencing people in their homes and sometimes in the workplace. In general, no attempts have been made to influence decisions made at city halls, in the boardrooms of multinational companies, or in the courts. Nevertheless, the bottom-up strategy of influencing people as consumers and citizens may profitably be complemented by attempts to understand people in specific, rare, and influential roles. This might help to bring about structural changes in society that would make it easier for everyone to behave in accordance with the needs and demands of the environment.

See Also the Following Articles

Attitude Measurement ■ Attitudes ■ Conservation Behavior ■ Environmental Design and Planning, Public Participation in ■ Restorative Environments

Further Reading

Ajzen, I. (2001). Nature and operation of attitudes. *Annual Review of Psychology*, *52*, 28–58.

Dawes, R. M. (1980). Social dilemmas. *Annual Review of Psychology*, *31*, 169–193.

De Young, R. (1993). Changing behavior and making it stick. The conceptualization and management of conservation behavior. *Environment and Behavior*, *25*, 485–505.

Geller, E. S. (2002). The challenge of increasing proenvironment behavior. In R. B. Bechtel, & A. Churchman (Eds.), *Handbook of environmental psychology* (pp. 525–540). New York: John Wiley.

Hardin, G. (1968). The tragedy of the commons. *Science*, *162*, 1243–1248.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decisions under risk. *Econometrica*, *47*, 262–291.

Manstead, A. S. R., & Parker, D. (1995). Evaluating and extending the theory of planned behavior. In W. Stroebe, & M. Hewstone (Eds.), *European review of social psychology* (Vol. 6, pp. 69–96). Chichester, UK: Wiley.

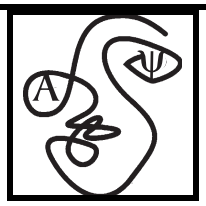
Schwartz, S. H. (1977). Normative influences on altruism. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 10, pp. 221–279). New York: Academic Press.

Staats, H. (2003). Understanding pro-environmental attitudes and behavior: An analysis and review of research based on the theory of planned behavior. In M. Bonnes, T. Lee, & M. Bonaiuto (Eds.), *Psychological theories for environmental issues* (pp. 171–201). Aldershot, UK: Ashgate.

Stern, P. C. (2000). Psychology and the science of human–environment interactions. *American Psychologist*, *55*, 523–530.

Stern, P. C., & Oskamp, S. (1987). Managing scarce environmental resources. In D. Stokols, & I. Altman (Eds.), *Handbook of environmental psychology* (Vol. 2, pp. 1043–1088). New York: John Wiley.

Verplanken, B., & Aarts, H. (1999). Habit, attitude, and planned behaviour: Is habit an empty construct or an interesting case of goal-directed automaticity? In W. Stroebe, & M. Hewstone (Eds.), *European review of social psychology* (Vol. 10, pp. 101–134). Chichester, UK: Wiley.



Prosocial Behavior, Development of

Nancy Eisenberg and Adrienne Sadovsky
Arizona State University, Tempe, Arizona, USA

1. Introduction
2. The Development of Prosocial Behavior
3. Gender Differences
4. Cultural Differences
5. Contextual Factors
6. Dispositional Factors
7. The Origins of Prosocial Behavior
Further Reading

GLOSSARY

altruism Prosocial behavior that is motivated by sympathy or moral values/concerns rather than by egoistic factors (e.g., concrete rewards, social approval, elimination of vicarious distress to make oneself feel better).

empathy An affective response based on another person's emotion or condition that is similar to what the other person is experiencing or would be expected to experience.

meta-analysis A procedure that statistically combines the findings from a number of different studies to determine whether a general finding holds across various empirical studies.

prosocial behavior Voluntary behavior that is intended to benefit another person.

sympathy Other-oriented concern that is often based on empathy.

Prosocial behavior starts to develop early in life and plays an important role in moral development. There are cultural and gender differences in some types of prosocial behavior, and prosocial actions are affected by contextual factors as well as by individual differences in temperament or personality. Prosocial tendencies are likely due to genetic factors as well as to socialization both within the family and in the child's larger social world.

1. INTRODUCTION

Prosocial behavior, or voluntary behavior intended to benefit another person, is of obvious importance for enhancing human relationships and for the smooth functioning of society. Of particular importance is altruism, which sometimes is defined as prosocial behavior that is motivated by sympathy or moral values/concerns rather than by egoistic factors (e.g., concrete rewards, social approval, elimination of vicarious distress to make oneself feel better). Unfortunately, it is often difficult to differentiate the two; indeed, some psychologists (and philosophers) claim that real altruism does not exist. Thus, most of the existing empirical work involves the study of prosocial behaviors that may or may not be altruistically motivated.

Some theorists, such as Batson, believe that altruism usually stems from an individual experiencing another's emotional state (i.e., empathy) or other-oriented concern (i.e., sympathy). Other theorists, such as Eisenberg, believe that prosocial behaviors (especially altruistic ones) can be engendered by internalized moral values

©2004 Elsevier Inc.
All rights reserved.

as well as by sympathy. Still others, such as Cialdini, believe that prosocial behavior is motivated by hedonism; for example, people help to improve their own moods. Empirical studies have demonstrated that there are associations between prosocial behaviors and both sympathy and measures of individuals' moral reasoning and values. Empathy and sympathy are also believed to inhibit cruelty toward others, although it is not clear whether prosocial behavior and aggression are highly negatively related (although both likely are affected by empathy and sympathy).

2. THE DEVELOPMENT OF PROSOCIAL BEHAVIOR

Even infants can experience concern for others and have the capacity to perform prosocial behaviors. By 10 to 14 months of age, infants frequently express agitation when viewing others in distress. Self-distressed reactions when viewing others in distress decrease in frequency during the second year of life, and by 18 months of age infants sometimes make efforts to comfort (e.g., by patting a person) or help other people who are upset or sad, especially their parents. In addition, even toddlers frequently try to assist adults with household chores.

According to Hoffman, toddlers do not fully distinguish between their own inner states and other people's inner states; consequently, their attempts to help are often egocentric (e.g., a toddler may bring a distressed peer to his or her own mother rather than the peer's mother). With age, children's prosocial actions become more situationally appropriate and sensitive to others' specific needs.

Meta-analytic reviews have shown that the frequency of children's prosocial behavior, including sharing, comforting, and instrumental helping, tends to increase with age across childhood, although there is little difference in the frequencies of prosocial behavior between younger and older adolescents. Age differences tend to be greater in studies with experimental/structured designs than in those with naturalistic or correlational designs, and they also tend to be greater when prosocial behavior is measured with self- or other-reports than when it is measured with observations. Moreover, although the consistency of prosocial behavior across contexts is low during the early years of life, it becomes more consistent with age and there are relatively clear individual differences in prosocial tendencies by late childhood and adolescence. However, there is some consistency of prosocial tendencies over time, even from the preschool years. For example, preschool children who spontaneously engage in sharing behaviors that involve some cost (e.g., giving up a toy) are more sympathetic and prosocial as adolescents and adults.

Younger and older children also differ in the reasons they report for their prosocial actions. Older children are more likely to say that they assist for apparently altruistic or other-oriented motives, whereas younger children are more likely to assist to obtain material or social rewards. Furthermore, with development, children are more likely to resolve hypothetical moral dilemmas regarding assisting others at a cost to themselves by referring to higher level moral justifications. For example, preschoolers tend to verbalize primarily hedonistic reasoning and some needs-oriented (primitive empathic) reasoning. By school age, some children begin to express concern about gaining approval, enhancing interpersonal relationships, and behaving in stereotypically "good" ways. During late elementary school or thereafter, children start to verbalize reasoning reflecting abstract principles, internalized affective reactions (e.g., guilt or positive affect about the consequences of their behavior for others, living up to internalized principles), self-reflective sympathy, and perspective taking. From childhood into early adolescence, children's needs-oriented reasoning increases, whereas hedonistic reasoning decreases. For most children, reasoning concerning role taking, internalized norms/rules/values, generalized reciprocity across members of society, internalized affective reactions based on concern about the consequences of their behavior for others, and positive affect related to internalized values and living up to those values is uncommon during elementary school and increases with age from adolescence into early adulthood. Nonetheless, people of all ages sometimes express low-level, hedonistic moral reasoning. Age-related changes in children's prosocial moral reasoning are of interest because higher level reasoning likely reflects developmental changes in the goals and motives underlying children's prosocial behavior and has been correlated with the tendency to engage in prosocial behavior.

3. GENDER DIFFERENCES

Gender stereotypes regarding sex differences in prosocial behavior are common in American culture and probably in numerous other cultures. In general, girls and women are expected to be more caring and to

engage in more prosocial behaviors than are boys and men. Consistent with the stereotype, in meta-analytic analyses, effect sizes indicate that girls reliably display greater prosocial tendencies than do boys. For example, in 1998, Eisenberg and Fabes found that girls evidenced considerably greater kindness and consideration for others than did boys (d = .42). Likewise, girls shared and donated more than did boys, although not to such a marked extent (d = .13), and also provided somewhat more instrumental help (d = .14). In additional meta-analyses, investigators found that girls exhibited greater empathy and sympathy than did boys, at least on some types of measures (e.g., self-reports, some observed behavioral measures). In 1999, Fabes and colleagues noted that these sex differences in prosocial tendencies become stronger during adolescence as compared with early and middle childhood. Thus, there is support for a developmental change in the expression of prosocial behaviors.

In contrast, examination of sex differences in adults' prosocial behaviors has provided mixed results. In 1986, Eagly and Crowley found that men, but not women, generally exhibited considerably greater helping (d = .34). However, many of the studies with adults involved helping strangers and performing instrumental tasks (e.g., changing a tire). These authors suggested that men's greater helping likely reflects the conventional gender role-rooted expectation for men to enact agentic behaviors or to "rescue" others (i.e., behave chivalrously) and that the findings likely differed for prosocial behaviors involving both caring and recipients who are known.

Consistent with Eagly and Crowley's speculations, women have been found to display greater prosocial tendencies than have men with regard to other indexes of prosociality. For example, women deliver messages containing greater emotional support and report more sympathy and empathy than do men, both on questionnaires and after being exposed to empathy-inducing stimuli, although they generally do not differ on physiological or facial measures of empathy. Furthermore, adolescent girls and adult women generally score higher on caring and other-oriented moral reasoning.

It is important to emphasize that the strength of sex differences fluctuated greatly across studies included in the aforementioned meta-analyses. Myriad factors, such as the design of the study, account for some degree of the difference between women and men. Nonetheless, the data do support the legitimacy of the stereotypes regarding sex differences in prosocial behaviors, at least when taking into account the type of helping behavior. However, the finding that sex differences in empathy and sympathy are stronger for self-reports than for other methods of assessment is consistent with the notion that people likely perceive a greater gender difference in prosocial tendencies than actually exists and respond accordingly.

4. CULTURAL DIFFERENCES

Prosocial tendencies appear to differ considerably across cultures. For example, in some traditional cultures such as the Ik in Africa, aggression and cruelty are commonplace, whereas in some other cultures, caring, helping, and supportive behaviors are fairly commonplace. Anthropologists have found that children in traditional cultures tend to assist family members and neighbors more than do children in industrial communities. Even in industrialized societies, children from collectivist cultures that emphasize the well-being of the larger social group (e.g., China) appear to be more prosocial with peers than do children from individualist cultures (e.g., the United States). However, it is not known whether these cross-cultural differences hold when unfamiliar people, rather than family members, neighbors, and friends, are the targets of aid.

The reasons for observed cross-cultural and subcultural differences in prosocial and cooperative behavior are probably multiple and complex. Clearly, the value placed on the well-being of others and on prosocial behavior varies across societies and likely is taught to children by parents and teachers as well as communicated through social norms. Some evidence suggests that these expectations are communicated in everyday life. For example, in cultures where children are expected to be cooperative and help others, often as part of their chores (e.g., when caring for siblings), children tend to be relatively prosocial. Moreover, prosocial actions tend to be viewed as more obligatory in collectivist societies than in individualist societies. Also in collectivist societies, privileges and social acknowledgment in the classroom are dependent on group accomplishments rather than individual accomplishments, so that children are reinforced for helping one another.

5. CONTEXTUAL FACTORS

Children's and adults' prosocial behaviors vary considerably as a function of the specific context. For

example, people are more likely to help if they see someone else—even a stranger—modeling prosocial behavior. Moreover, there is evidence that children and adults help more when they are in a good mood and help less when they are sad or depressed (especially in the case of older children and adults). Furthermore, factors such as the clarity of cues related to the other person's need, the cost and benefits associated with a particular prosocial action, and the salience of social norms regarding prosocial behavior appear to influence whether people choose to help or share with others.

6. DISPOSITIONAL FACTORS

Some people are more prosocial than others. Individual differences in prosocial behavior have been linked to a variety of aspects of personality or other dispositional characteristics. For example, as already noted, people who are prosocial are relatively prone to experience empathy and sympathy for others. In addition, prosocial individuals tend to feel a sense of responsibility for their actions and for others and also feel competent to assist. During childhood, prosocial tendencies also have been linked to being well regulated, socially competent, emotionally positive, and sociable, although nonassertive and less sociable children may be especially likely to perform compliant prosocial behaviors that are requested by their peers (in part because they have difficulty in asserting their own will). Thus, even within cultures and across the sexes, characteristics of individuals seem to contribute to their willingness and ability to assist others.

7. THE ORIGINS OF PROSOCIAL BEHAVIOR

Findings in twin studies support the common assumption that there is a genetic component to prosocial behavior and the tendency to empathize with others. Nonetheless, individual differences in children's socialization experiences in the home, school, and community are linked to individual differences in their prosocial behavior.

Prosocial children (and adults) tend to be those with parents who model and value prosocial actions and who apply ethical principles (especially those related to caring and helping) to a broad spectrum of humanity. Parents of prosocial children also tend to use reasoning (inductions) when they discipline their children (e.g., pointing out the consequences of children's behavior for others) and seldom resort to physical punishment or threats of punishment. In addition, these parents tend to be authoritative and democratic in their parenting style (i.e., are supportive and use inductions while also clearly communicating their expectations for appropriate behavior) and provide opportunities for their children to assist others. In general, parents of prosocial children also appear to promote the development of children's sympathy, and sympathy appears to motivate altruistic caring behavior.

Similarly, prosocial and cooperative behavior also can be fostered in school-based programs that emphasize rational discipline, cooperation, and prosocial values and that include activities designed to heighten children's sympathy and understanding of others and their feelings. In addition, a positive relationship between children and their teachers is associated with children's consideration for others and children's empathy with unfamiliar peers. Finally, there is a modest positive relation between children's viewing of prosocial television programming and their prosocial behavior. Thus, it appears that individual differences in prosocial behavior are due to a variety of influences in the home, media, and community.

See Also the Following Articles

Emotion ■ Gender and Culture ■ Stereotypes

Further Reading

Batson, C. D. (1991). *The altruism question: Toward a social–psychological answer*. Hillsdale, NJ: Lawrence Erlbaum.

Davis, M. H. (1994). *Empathy: A social psychological approach*. Madison, WI: Brown & Benchmark.

Eagly, A. H., & Crowley, M. (1986). Gender and helping behavior: A meta-analytic review of the social psychological literature. *Psychological Bulletin*, *100*, 283–308.

Eisenberg, N., & Fabes, R. A. (1998). Prosocial development. In N. Eisenberg (Ed.), *Handbook of child psychology,* Vol. 3: *Social, emotional, and personality development* (5th ed., pp. 701–778). New York: John Wiley.

Eisenberg, N., & Mussen, P. (1989). *The roots of prosocial behavior in children*. Cambridge, UK: Cambridge University Press.

Fabes, R. A., Carlo, G., Kupanoff, K., & Laible, D. (1999). Early adolescence and prosocial/moral behavior: I. The role of individual processes. *Journal of Early Adolescence*, *19*, 5–16.

Hoffman, M. L. (2000). *Empathy and moral development: Implications for caring and justice*. Cambridge, UK: Cambridge University Press.

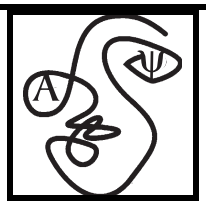
Jaffee, S., & Hyde, J. S. (2000). Gender differences in moral orientation: A meta-analysis. *Psychological Bulletin*, *126*, 703–726.

Oliner, S. P., & Oliner, P. M. (1988). *The altruistic personality: Rescuers of Jews in Nazi Europe*. New York: Free Press.

Schroeder, D. A., Penner, L. A., Dovidio, J. F., & Piliavin, J. A. (1995). *The psychology of helping and altruism: Problems and puzzles*. New York: McGraw–Hill.

Whiting, B. B., & Whiting, J. W. M. (1975). *Children of six cultures: A psychocultural analysis*. Cambridge, MA: Harvard University Press.

Zahn-Waxler, C., Robinson, J., & Emde, R. N. (1992). The development of empathy in twins. *Developmental Psychology*, *28*, 1038–1047.



Psychological Assessment, Standards and Guidelines for

Dave Bartram

SHL Group, Surrey, United Kingdom

1. Introduction
2. International Standards and Guidelines
3. National Standards and Guidelines
4. Conclusion
   Further Reading

GLOSSARY

psychological assessment Any assessment of an individual's psychological characteristics carried out using procedures that are standardized and repeatable; this can include structured interviews and standardized work sample assessment exercises as well as more traditional psychological tests.

psychological testing Standardized procedures whereby quantitative measures of psychological attributes or characteristics are obtained.

psychological tests Instruments with known psychometric properties designed to measure one or more psychological characteristics; the psychometric properties will include measures of reliability and validity as well as data to aid interpretation.

standards and guidelines Two kinds of criteria that differ in that standards are generally more prescriptive than guidelines, whereas guidelines tend to be advisory and informative.

test user Anyone who is responsible for the use of a psychological test in an assessment process.

Standards and guidelines in assessment cover three related areas: the people who use tests, the tests the people use, and the process of assessment within which tests are used. Both nationally and internationally, all three of these issues have been addressed in various ways.

1. INTRODUCTION

Standards and guidelines have been developed to cover all aspects of test design, development, and use. In many countries, test publishers or distributors were the first to develop qualification criteria, guidelines, and standards to regulate, advise, and control test users. National psychological associations and international bodies concerned with testing and test use (e.g., European Federation of Psychologists Associations [EFPA], International Test Commission [ITC]) have also played a major role in developing and promulgating standards and guidelines. These standards and guidelines have addressed issues associated with test development, the qualification of test users and test distribution, proper test use and best practice, and test taker rights and responsibilities.

Standards and guidelines cover all sorts of psychological testing and assessment methods—from tests of ability and aptitude to tests of personality and motivation, from paper-and-pencil tests to computer-based

©2004 Elsevier Inc.
All rights reserved.

and Internet-delivered tests, from single instruments to complex 2- or 3-day "assessment centers."

A survey of 48 countries, carried out jointly in 1996 by the ITC and EFPA, suggested the following:

- Approximately 79% of test users operate in the educational field, 12% in the clinical field, 9% in work and organizational settings, and less than 1% in the forensic area.
- The majority (86%) of test users are nonpsychologists.
- The clinical and forensic areas are the only ones where nonpsychologists do not outnumber psychologists as test users.
- Only 41% of all users have received any specific training in test use, with the lowest rates of training in the educational area and the highest rates in the clinical area.
- Psychologists have only slightly higher rates of training than do nonpsychologists.

Overall, the survey showed the need for more training for all test users and revealed marked differences in patterns of response among different countries. Of particular interest were the big differences among countries in the power and role of the professional psychological associations in relation to setting and enforcing standards.

2. INTERNATIONAL STANDARDS AND GUIDELINES

In the relatively recent past, it was possible to think of individual countries as "closed systems." Changes could be made in terms of best practice, procedures, and laws affecting the use of tests in one country without there being any real impact on practice or laws in other countries. People tended to confine their practice of assessment to one country, and test suppliers tended to operate within local markets by adapting prices and supply conditions to the local professional and commercial environments.

This situation has changed. Individual countries are no longer closed systems. Psychological testing and assessment is an international business. Many test publishers are international organizations that sell their tests in a large number of countries. Many of the organizations using tests for selection and assessment at work are multinationals. In each country, test suppliers and test users are likely to find differences in practices related to access, user qualification, legal constraints on test use, and so on.

2.1. The International Test Commission

The ITC is an association of national psychological associations, test commissions, publishers, and other organizations committed to promoting effective testing and assessment policies and to ensuring the proper development, evaluation, and use of educational and psychological instruments. The ITC has been responsible for a number of guideline development projects. In each case, the process has been one of ensuring widespread international agreement on the content of the guidelines.

2.1.1. ITC Guidelines on Adapting Tests

These provide 22 guidelines for adapting psychological and educational tests for use in various different linguistic and cultural contexts. These guidelines fall into four main categories: the cultural context, the technicalities of instrument development and adaptation, test administration, and documentation and interpretation. Adaptation is an area of major importance as tests become used in more and more countries and as tests developed in one country get translated or adapted for use in another country. Adaptation needs to consider the whole cultural context within which a test is to be used.

2.1.2. ITC Guidelines on Test Use

These guidelines were published in 2000 and have been widely adopted by countries around the world as the basis for their own local standards. They focus on good test use and cover both ethical issues and best practice process. They provide a common international framework from which specific standards, codes of practice, qualifications, user registration criteria, and so on can be developed to meet local needs. By 2004, these guidelines were available in 14 different languages.

The competencies defined by the guidelines are specified in terms of assessable performance criteria. These competencies cover the following:

- Professional and ethical standards in testing
- Choice and evaluation of alternative tests
- Test administration, scoring, and interpretation
- Report writing and feedback
- Rights of the test candidate and other parties involved in the testing process

2.1.3. ITC Guidelines for Computer-Based and Internet-Delivered Testing

Work is in progress on the development of international guidelines in this area. These were scheduled for publication in 2004.

2.2. European Initiatives

2.2.1. The EFPA Test Review Criteria

The EFPA represents 30 European countries. Its Committee on Tests and Testing published the EFPA standards for the review of psychological tests in 2001. These standards, developed from British and Dutch work on test reviewing, are now being used as the basis for test reviewing across Europe and provide the basis for a test registration procedure.

Within Europe, new legislation on data privacy has had a major impact on testing and the use of tests. This has resulted in good practice in assessment becoming embodied in law and has led to the development of guidance to support this within European countries. Increasingly, it is being recognized that there is a need for standards and guidelines to be considered at a European level to match the demands arising from European Union directives, especially those that relate to fairness and equity issues in the employment area and to the transferability of professional qualifications among countries.

2.2.2. Other European Initiatives

The European Association of Psychological Assessment (EAPA) has published a set of Guidelines for the Assessment Process. The European Association of Work and Organizational Psychologists (EAWOP) has established a working group to focus on Europe-wide issues relating to testing and assessment in the area of work and employment. Work has started on the development of European standards for the certification of competence in the use of tests for assessment in the work and organizational field. This work is being carried out by a joint EFPA–EAWOP task force and could lead to a European qualification of competence in testing.

2.2.3. Other International Initiatives

International standards and guidelines have been developed in some other areas that relate to psychological testing and assessment. For example, an international task force of leading practitioners in the field developed Guidelines and Ethical Considerations for Assessment Center Operations. These guidelines have been updated a number of times.

3. NATIONAL STANDARDS AND GUIDELINES

International guidelines and standards have both influenced national initiatives and been influence by them.

3.1. The United States

There are a large number of organizations and committees involved in work that relates to testing and assessment practice in the United States. It is not possible to mention all of these here. However, some key ones are described in what follows.

3.1.1. The "Joint Standards"

The Standards for Educational and Psychological Testing (the 1985 version of the standards issued jointly by the American Educational Research Association [AERA], the American Psychological Association [APA], and the National Council on Measurement in Education [NCME]) have been widely adopted around the world as the authoritative definition of technical test standards. After a lengthy revision and consultation process, a new version of these influential standards was published in 1999. Many countries' psychological associations were expected to adopt these as the successor to the earlier edition.

3.1.2. Test User Qualification

Over the past few years, there has been some progress in the United States on the issue of test user qualification. In 2000, the APA Task Force on Test User Qualifications (TFTUQ) produced guidelines that inform test users and the general public of the qualifications that the APA considers to be important for the competent and responsible use of psychological tests. In this context, the term "test user qualification" is being used to indicate a desired level of competence rather than the award of some certificate or license or the outcome of a credentialing process (as is the case in the United Kingdom). The report suggests that the old three-level APA system (A, B, C) for classifying test user qualifications that was first defined in 1950 is now

obsolete and needs to be replaced by one where the competences required are defined in relation to the context, instrument, and use to which it will be put.

3.1.3. Computer-Based Testing

The Association of Test Publishers (ATP) has produced technology-based guidelines for the testing industry. The goal of the proposed guidelines is the international adoption by companies involved in technology-based testing. The guidelines are for use as an aid in the development, delivery, and use of computer-based certification examinations as well as psychological testing in general. The guidelines cover applications for use on the Internet and various multimedia computer strategies used to deliver, administer, and score tests.

3.1.4. Internet-Delivered Tests

In 2001, the APA set up a task force on psychological testing and the Internet to assess the extent and range of testing on the Internet and to review current practices. It aims to define and outline the issues raised by Internet testing, including test validity, administration, confidentiality of test taker and test results, test taker authenticity, ethical interpretations of test results (e.g., feedback), psychological dynamics of Internet testing (e.g., proclivity of being more revealing when taking Internet-based tests), copyright infringement, psychometric equivalence (e.g., comparability of tests results), license issue of the psychologist (e.g., crossing state lines), and interpretations based on limited assessment information.

3.2. European Countries

There is a great deal of work going on in Europe in relation to the development and encouragement of best practice in psychological assessment. Some major examples are described in what follows.

3.2.1. The United Kingdom

The United Kingdom was the first country to establish a test user certification procedure accredited by the national psychological association, the British Psychological Society (BPS). Both psychologists and nonpsychologists can obtain certificates in test use from the BPS. These qualifications are subject to stringent quality control or "verification" procedures and currently provide recognition of qualification in test use relating to occupational assessment. Similar qualifications for educational testing became available in 2004. All of these qualifications are based on a set of performance standards defining the competence required of a test user and a set of guidelines defining how this competence should be assessed.

The BPS has a standardized process for the review of tests and has published a number of volumes of test reviews focused on assessment in the work and organizational assessment area. These are designed as "consumer" reviews for test users rather than being for a more academic audience.

All of the BPS procedures and services relating to testing and test use are now managed within a single organization, the Psychological Testing Center. This organization manages the delivery of test user certification and test reviewing and operates a Web site that provides information, advice, and guidance for test developers, test users, and test takers.

The BPS has produced a set of guidelines for assessment centers and standards for computer-based testing. The BPS has also produced guidelines on 360-degree feedback in collaboration with the Chartered Institute for Personnel and Development, the University of Surrey at Roehampton, the SHL Group, and the Department for Trade and Industry.

The British Standards Institute has issued a standard (BS7988) covering technology-based assessment. This standard focuses on many of the technical issues involved in the delivery of assessments by computer and over the Internet and is concerned mainly with educational testing and tests for licensing and certification. Unlike the ATP guidelines, this standard currently excludes areas of psychological testing (e.g., personality assessment).

3.2.2. Germany

In Germany, the national standards institute has published a "norm" setting out standards for assessment for job selection (DIN33430). These standards set out good practice for the process of assessment and include requirements relating to test quality and test user competence.

3.2.3. Finland

Guidance on good practice in assessment has been produced in Finland as a consequence of recent legislation relating to the European Union Directive on Data Privacy. This directive has had implications for all European countries on how test data can be used, the rights of test takers over their data, and what is done with the data.

Like the German standard, the Finnish guidance describes best practice for assessment process as a whole (in employment settings) rather than just for psychological testing. Recently, the Finnish Psychological Association developed test user certification procedures for psychologists and nonpsychologists based on this guidance.

3.2.4. Sweden, Norway, and Denmark

In Sweden, an institute, Stiftelsen för Tillämpad Psykologi (STP), was set up in 1996 to provide test user qualification certification, test reviewing, the quality audit of organizational testing policies, and an ombudsman function to which test takers may appeal, Test and test use quality assurance procedures build on the ITC Guidelines on Test Use and on the BPS test review and test user certification models. The STP also accredits organizational test policies and testing processes. Guidelines for organizational policies on test use are based on the model provided in the ITC Guidelines on Test Use.

More recently, Det Norske Veritas (DNV) has initiated a project to establish test user certification procedures in Norway, Sweden, and Denmark. DNV, one of the leading worldwide registrars for certification of ISO 9000 and related standards, is working in conjunction with the Norwegian Psychological Association, the STP, and representation from Denmark.

3.2.5. The Netherlands

In the Netherlands, Commissie Testaangelegenheden Nederland (CoTAN), a committee of the Dutch Psychological Association, has a long history of producing reviews of tests covering all areas of assessment.

More recently, the Institute for Best Test Practice (4TP) was set up with a set of aims and objectives very similar to those of the Swedish STP. The Dutch institute aims to improve the quality of test use in occupational and workplace assessment. It is making use of the ITC guidelines and the experience of both the BPS and the STP in setting standards with which users, instruments, and practices must confirm. The institute has developed procedures for examining the competence of test users, for reviewing tests (using the EFPA criteria), and for auditing organizations.

A guiding principle of 4TP is the need for transparency. First, the standards adopted must be open and available for all to see. Second, all of the results of examinations carried out by the institute will be published on the Internet—registers of certified test users, reviewed tests, and quality audited organizations. Therefore, everybody interested in the quality of people, instruments, and practices will be able to access this information. In this way, it is hoped that the institute will provide a practical way of implementing quality and transparency.

4. CONCLUSION

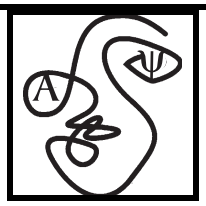
The challenge for the future is to bring together the various standards and guidelines from different countries and organizations and to secure international agreement on the key issues. Work by bodies such as the ITC, EFPA, and EAWOP has made considerable progress during the past few years in moving this forward.

See Also the Following Articles

Intelligence Assessment ■ Personality Assessment ■ Psychometric Tests ■ Psychophysiological Assessment

Further Reading

American Psychological Association. (n.d.). www.apa.org.
Association of Test Publishers. (n.d.). www.testpublishers.org.
Bartram, D. (1996). Test qualifications and test use in the UK: The competence approach. *European Journal of Psychological Assessment*, *12*, 62–71.
Bartram, D. (1998). The need for international guidelines on standards for test use: A review of European and International initiatives. *European Psychologist*, *3*, 155–163.
Bartram, D. (2001). The development of the International Guidelines on Test Use: The International Test Commission Project. *International Journal of Testing*, *1*, 33–54.
Bartram, D., & Coyne, I. (1998). Variations in national patterns of testing and test use. *European Journal of Psychological Assessment*, *14*, 249–260.
British Psychological Society. (n.d.). www.psychtesting.org.uk.
Buros Institute. (n.d.). www.unl.edu/buros/.
European Federation of Psychologists Associations. (n.d.). www.efpa.be.
Honkanen, H., & Nyman, K. (2002). *Handbook of good practice in personnel assessment*. Helsinki: Finnish Psychological Association.
International Test Commission. (n.d.). www.intestcom.org.
Institute for Best Test Practice. (n.d.). www.4tp.nl.



Psychological Rehabilitation Therapies

Karen G. Langer

New York University Medical Center, New York, New York, USA

1. History
2. Populations Served
3. Treatment Considerations
4. Psychological Issues
5. Therapeutic Approaches
6. Conclusion
   Further Reading

and/or self-sufficiency have been partially or extensively compromised temporarily or long term; interventions are aimed at restoring or improving level of independence or self-sufficiency, enhancing personal functioning and quality of life, or addressing the physical, architectural, social, cognitive, communicative, and/or other barriers to everyday life.

GLOSSARY

acquired disability A loss, impairment, or decline from the previous level of functioning or bodily integrity, sustained due to illness, trauma, or injury.

brain dysfunction A brain disturbance that can be caused by illness or injury and that results in disruption to life-sustaining autonomic functions or to motor, sensory, appetitive, affective/emotional, and/or cognitive functions subserved by the brain areas involved.

cognitive remediation Provision of treatment strategies geared to improve cognitive abilities directly or to teach compensation strategies for cognitive skills weakened by brain dysfunction and help people deal with changes in cognitive functions.

congenital disability A functional or physical impairment originating from a developmental event or deviation during or prior to birth of nonhereditary origin.

rehabilitation An interdisciplinary treatment approach, devoted to the optimization of physical and psychological functioning, where independence, personal autonomy,

Psychotherapy has among its goals establishing a new, or restoring a prior, sense of identity to patients whose sense of self has been deeply threatened by mental or physical trauma. Along with decreasing emotional suffering, the restoration of identity may be a key outcome of psychotherapeutic treatment in rehabilitation. Patients' losses include not only the physical aspects of disability but also the symbolic and psychological losses that ensue. Cognitive deficits may develop as a result of brain dysfunction, and these may also be targets of remediation techniques designed to compensate for, or directly strengthen, those functions that were weakened. Emotional sequelae, including depression, anxiety, terror, anger, blunting, constriction, instability, lability, and agitation, may also be foci of treatment. Broader issues involving psychosocial and interpersonal needs, physical dependence, personal care routines, medical comorbidities, vocational disruptions, and financial losses all can have impact. All that forms the fabric of quality of life may need to be

©2004 Elsevier Inc.
All rights reserved.

considered in treating patients with disabling loss in rehabilitation.

1. HISTORY

Among the early developments in the field of rehabilitation were concerns for the holistic treatment of people with physical disabilities. Dr. Howard Rusk, the pioneer of rehabilitation medicine as a specialty within medical practice, was dedicated to the notion that persons served were best viewed as people within a complex of not only physical but also psychological and social facets. The latter would include a focus on the emotional effects of disability and the social implications as well as on the vocational and avocational concerns that need to be considered.

The field of rehabilitation psychology was an outgrowth of the field of rehabilitation medicine in the sense that psychologists devoted themselves to the study and treatment of persons with acquired or congenital disabilities. Psychological treatment and psychotherapeutic treatment of persons with disabilities are now considered to be an integral part of rehabilitation medicine and behavioral medicine programs. Psychological rehabilitation draws heavily from fields such as rehabilitation, clinical psychology, neuropsychology, health psychology, behavioral medicine, counseling psychology, and developmental psychology. Practitioners come to the field from one of these areas and gain the specialized training necessary to confront the unique challenges encountered with rehabilitation of patients in psychological treatment.

2. POPULATIONS SERVED

Psychologists serve in adult and pediatric settings that include all ages of the life span—pediatric, adolescent, general adult, and geriatric. The populations served in rehabilitation and medical settings involve a great diversity of conditions. These include traumatic injuries of all kinds and illness-related disabilities such as neurological disorders, neuromuscular disorders, cancers, debility due to deconditioning from prolonged illnesses, amputations, spinal cord injuries, brain dysfunction, head traumas, and orthopedic conditions—in short, any conditions that may result in physical limitations or disabilities. Patients may have sustained multiple injuries in ski accidents or falls or may have been involved in car or diving accidents, resulting in spinal cord or head injuries. They may have been victims of natural disasters or terrorist attacks, collapsed buildings, gunshot injuries, burns due to fires or explosions, or other tragic events. Patients may have undergone back surgery for spinal conditions such as stenosis or may have undergone hip replacement surgery due to osteoarthritis or degenerative joint disease. Others may have more chronic conditions such as Parkinson's disease, lupus, and multiple sclerosis.

Onset may be sudden such as in stroke, myocardial infarction, and Guillain–Barre syndrome. Alternatively, sudden onset (e.g., lower extremity amputation) can sometimes follow a prolonged period of declining function/disability and pain or even progressive surgeries. Time factors may play a role in patients' experience of disability, including chronicity and duration of condition, age at onset, and congenital versus acquired disability. For some patients, the longer the condition has existed, the more time they have had to adjust; for others, the pain or disability may have persisted for a long time, so that the stress can be additive or cumulative. Other factors to be considered include the nature of onset (e.g., traumatic vs nontraumatic, accident vs illness related), risk-to-life factors (e.g., initial brush with death, terminal nature of condition), and prognostic descriptors (e.g., progressive, stable, remitting, expected to improve). Also, the treating clinician must consider the question of awareness, that is, what the patient has been told, whether the patient knows the condition, and the patient's ability to be aware of implications based on emotional factors such as denial.

Psychological treatment in rehabilitation often deals not only with the stresses imposed by disability and altered life circumstances but also with issues of hospitalization and medical tests and procedures. Posttraumatic stress disorder may be an aspect of patients' emotional responses as they relive the accidents or traumas, sometimes in the form of flashbacks or nightmares. Naturally in the course of continuing treatment, psychologists may deal with a broad range of evolving and changing responses to the whole spectrum of effects of disability and its life-altering impact.

3. TREATMENT CONSIDERATIONS

3.1. Assessment

Psychological intervention often follows an assessment that can range from 1 hour to several hours and

generally taps all pertinent aspects of emotional response to the loss, change, or trauma. Emotional reactions may include anger, anxiety, depression, agitation, emotional lability, emotional stability, and even apparent indifference. Signs of affective blunting or constriction may be noted and need to be distinguished from illness-specific features such as the mask-like face (i.e., masked facies) often associated with Parkinson's disease and prosodic deficits (i.e., aprosodia) often associated with right hemisphere brain dysfunction. Included in this assessment is a focus on psychosocial history, vocational history, health habits (e.g., dietary, exercise, smoking, alcohol, recreational drugs), medical history, premorbid psychological functioning and history of treatment, suicide risk, ability to communicate emotional needs based on affective constriction and defensive styles, awareness of losses, emotional denial, cognitive ability to understand and appreciate the losses, prior risk-taking behaviors, educational history, and living arrangements. Also tapped are features of mental status and cognitive function as indicated through cognitive or neuropsychological testing during the process of psychological assessment.

3.2. Setting

The setting in which the patient is treated is also meaningful. Early phases of rehabilitation, which are often acute care or hospital based, may engender needs that are centered around medical or nursing needs, personal care routines, bodily functions, and/or food intake. Treatment often involves psychoeducation along with cognitive–behavioral or psychodynamic approaches to help educate a frightened patient who is unfamiliar with medical/hospital routines and terminology. Issues of dependence on caregivers—often strangers—for basic personal needs can engender feelings of helplessness, shame, anger, and/or depression as the patient grapples with losses to dignity, self-esteem, and sense of control over bodily functions. The treating psychologist must be familiar with the physical status and needs of the patient. By conveying knowledge of the medical/physical needs and a sense of competence, the experience becomes more normalized for the patient, and expectations are addressed from both educational and psychological perspectives.

3.3. Nature of Stressors

Disability itself produces stressors that are cumulative or interactive. Physical losses, including motor deficits and sensory losses, affect all aspects of mobility, elimination functions, and activities of daily living. Speech or swallowing deficits may ensue, affecting quality of life in terms of communication and basic appetitive pleasure. Personality changes or cognitive losses may have far-reaching effects on sense of self, interactions with family and others, and ability to resume prior activities such as work and leisure activities. The Americans with Disabilities Act of 1990 offers guidelines and legal protection to people with disabilities regarding access to public services, transportation, employment, and so forth. Yet the modifications and the special equipment needs, as well as their visibility, may also be stressful. The fact is that a physical disability can involve disruption not only to the body and its functions but also to more psychological entities. In fact, among the most devastating losses experienced by the patient may be losses to aspects of the self. The role of cultural factors and attitudes is also critical in shaping the individual's responses to a disability and the responses of others in that subculture. Nonetheless, familial and individual differences are many and may have even more meaning to the patient than to the subculture alone.

3.4. Developmental Considerations

For congenital conditions producing physical or cognitive disabilities, there are complex issues pertaining to the development of sense of self involving body image and sense of mastery. These may be particularly salient for the child who is growing up with a physical disability. Patients may grapple with guilt regarding the impact of their disabilities on their families. In some instances, especially if there are genetic components to a disabling illness, there may be added stress on the parents' marriage and concerns for genetic transmission. Some adults with lifelong disabilities have come to develop special skills that help them to adapt in the world, including social skills, navigational skills, and personality features. These skills, which are highly prized for their adaptive value, may form a core part of the identity in a very positive way. Identification with disability may then be very strong and complex. Patients with exacerbation of long-standing disabilities (e.g., postpolio syndrome) often find it difficult to take advice from less seasoned staff members who might not know their conditions as intimately as do the patients.

Other developmental considerations pertain to stage of life, such as when the disability occurred during the

life span, prior experience with illness or disability in oneself or others, and the availability of supports. The latter is crucial in working with an aged population where many family members, who could be of support, may have already died or become disabled themselves.

4. PSYCHOLOGICAL ISSUES

The issues that are confronted by the patient and psychotherapist are many and include mortality, mutilation, loss of limb, helplessness, loss of body functions, alterations in sexual functioning, loss of psychological entities such as sense of competence and mastery, loss of personal identity and ways of knowing oneself that may be shattered or threatened by physical injury, loss of job, loss of one's sense of manhood or womanhood, and role alteration within the family or with friends. Patients are frequently frightened of repeating trauma and may have flashbacks or dreams of impending disaster anticipated with devastating illness. In end-stage illness, pain management is a concern, as is worry about losing one's mental faculties, whether due to the illness or due to medication. Sometimes, there are issues of unawareness, and these are among the most challenging ones because the patient might not even realize the severity of the situation. This unawareness is often due to cognitive dysfunction itself but is sometimes due to significant emotional denial that blocks the awareness and the ability to make sound judgments regarding care or planning needs. There may be overestimation of abilities that can pose ethical challenges to the treating therapist with regard to issues of competence. These often do not require legal/judicial intervention because there might also be the possibility of improvement in the patient's status.

Psychological treatment of the patient in rehabilitation is aimed at the emotional sequelae of the event onset or ongoing situation of loss. Reactions such as depression, anxiety, anger, agitation, emotional lability, and instability of mood are very common. The supportive normalization of certain reactions is a key aspect of forming the therapeutic bond and educating patients, particularly less medically sophisticated patients, about what to expect. For many others, a more aggressive approach to treatment of the significant psychic distress and emotional disorders is indicated. The notion of losing control not only over physical functions and body parts but also over the emotional response system can be very threatening for patients who previously took pride in emotional coolness. Patients who used denial or avoidance as defensive styles prior to the onset of their illnesses/disabilities will continue to use these defenses in dealing with the disabilities. Denial, as a response to disability, poses a significant challenge to treatment not only by psychotherapists but also by other treatment providers such as physicians, rehabilitation staff, and family members or caregiving aides.

The degree to which there is distress and suffering is crucial to consider in any psychological treatment. A depressive reaction may be chronic whereby the disability and its privations persist over years or decades. Facilitating adaptation to disability may then be a realistic goal for the psychologist. Adaptation, unlike "acceptance," means that an individual is coping with and dealing with the situation to the best of his or her abilities. Adaptation does not imply that the person has accepted a notion of permanence or resignation, and it certainly does not imply that the person has lost hope of improvement (although he or she might very well have). Restoring identity to the patient whose physical or cognitive losses have compromised his or her identity is an integral part of rehabilitation and therapy.

5. THERAPEUTIC APPROACHES

Psychological treatment may evolve from crisis intervention to approaches that address deeper issues as the patient emerges from the acute phase of shock or traumatic stress. In time, therapies using more depth or psychodynamic models may be chosen to work on emotional issues, or cognitive behavioral approaches may be used to deal with depression, anxiety, or other emotional responses. Stress management techniques are employed to help relieve anxiety, tension, and posttraumatic stress. Other models, such as existential explorations of meaning of loss and life, can be selected for use by the therapist as indicated.

Once adapted to the needs of the setting and the patient, traditional psychodynamic approaches are helpful to the patient in dealing with new or unresolved needs and conflicts that may pertain directly to the disability. Issues regarding dependence and care, having developmental roots, or reflecting challenges to the personality structure will often surface as the patient grapples with unwelcome changes in physical functioning. The impact on well-established patterns of interactions and roles within family and other social structures assume primary importance to the

patient. Thus, psychodynamic approaches can be very useful in helping patients, as well as staff and family, to understand the basic emotional features that are threatened by the disability (e.g., the need for self-reliance or control) or are exacerbated by the disability (e.g., fears about abandonment or loss of love). Historical roots of needs or conflicts that have remained dormant prior to the onset of disability may be more amenable to exploration since the onset and may allow for unexpected resolution. Support, empathy, and genuine caring can be integral parts of the healing process in adjusting to the trauma of loss and disability and of the conflict/need resolution process in using the psychodynamic approach to treat the patient in rehabilitation.

Stress management and relaxation techniques, alone or in conjunction with the other forms of psychological treatments offered, are very useful in the treatment of the medical or rehabilitation patient.

Other psychological treatments may involve behavioral management strategies employed when behavioral disruptions occur. These disruptions can be induced by a stress reaction in the context of usually more limited coping skills. Alternatively, they can occur as a result of restricted access to the patient's previously used coping skills due to the limitations of disability itself. There may also be underlying psychiatric disturbance that is exacerbated or aggravated by the psychic trauma of disabling losses, leading to psychic regression or decompensation. Sometimes, there may be primary effects of brain dysfunction on behavior (e.g., initiation, sexual disinhibition, problem-solving skills, memory, attention/concentration, impulsivity, agitation, aggressive responses, adynamia, apathy). These effects may disrupt behavior to a greater or lesser degree, and when they are seen in the social sphere or with family, they can be very disturbing. Often, the psychologist treats the patient's behavior through a combination of psychotherapy and behavioral conditioning principles. Behavioral management meetings may be organized to help staff implement individualized approaches to behavioral management issues, whether due to significant cognitive dysfunction or the patient acting out under stress.

Because some patients might not even realize a need for psychological treatment, they may appear to be uncooperative, defensive, unmotivated, and/or agitated. Judgment and safety may be affected by lack of awareness or by impulsive behavior. Patients might not even fully realize the implications of their condition, or may not yet realize the extent of the change in their behavior or emotional responses. These are challenges to treatment and may put the burden of goal setting on the treating therapist until these patients can better acknowledge changes and appropriate treatment goals.

With regard to the cognitive impairments that may accompany certain disabilities, specifically those involving brain function of various etiologies, there is a tradition of cognitive remediation treatments. Such treatments are frequently adjunctive to treatments offered by other disciplines, such as occupational therapy and speech therapy, and are often integrated into the patient's physical therapy program. Therapy is tailored to the nature of the deficit, so the patient may be treated for basic disturbances. Deficits are measured by using neuropsychological assessment tools if possible. If deficits are too severe, they are measured by observational assessment and by comparison with scales of recovery after brain injury. Neuropsychological deficits may involve arousal level (i.e., hyperaroused or hypoaroused), attention and concentration, orientation, thought coherence, visual perception, visuospatial skills and visual attention, reasoning, coherence, judgment, planning, initiation, response latency, organization, inhibition of response and problem solving, all aspects and types of memory, and language functions (with the latter being treated primarily by speech and language pathologists), among others. Basic functions, such as reading, writing, copying, and arithmetical skills, may be affected by any of these deficits.

For brain dysfunction, model systems have been designed to integrate the cognitive–behavioral treatments within the entire rehabilitation treatment program throughout the day. Various brain injury facilities have established treatment programs that integrate the conceptualizations of neuropsychological recovery after brain injury. These programs may exist ideally within a dedicated unit, but they also may be effective in "units without walls." These are designed to have the complete treating staff integrate the recommended cognitive approaches within their disciplines, tailored to the individual cognitive, behavioral, physical, and emotional losses of the patient with diffuse or traumatic brain injury.

Cognitive remediation treatment may target skills lost or weakened by brain dysfunction and may involve either direct remediation of the weakened skill or compensatory strategies that help the patient deal with and compensate for the losses (e.g., training in the proper use and organization of a log/notebook for the patient with memory storage/retrieval difficulties). There is some debate in the literature regarding the

efficacy of compensatory versus direct remediation strategies. Both individual remediation and group remediation have been used in different settings. Psychological treatments have also been used in novel ways to enhance physical functioning (e.g., biofeedback techniques to improve aspects of motor control in certain patients).

Treatment of comorbid psychiatric or substance abuse issues may pose a challenge in the medical rehabilitation setting. However, these issues may arise with a patient who has a preexisting personality disorder (with aggressive anger, impulsivity, risk-taking, and/or acting-out behaviors). In fact, the personality trait may perhaps be expressed in more exaggerated form due to the emotional stress of loss or to the direct or secondary effects of brain injury. Sometimes, particularly in the aftercare or outpatient setting, the patient may even be referred to the psychologist specifically for treatment of these disturbances (e.g., impulsivity, anger/aggression management). The impact on significant others and family may be profound, especially if there is a change or worsening of a preexisting problem or a new emergence of a disturbing behavioral symptom. Family members may feel at a loss to deal with these disruptions, and counseling may be beneficial.

6. CONCLUSION

Depending on the setting, training, and need, therapists may take a more active or directive role. Some may become involved in coordination on behalf of patients for social services and/or vocational needs; with agencies, the challenge is to help the patients rebuild their lives. Others may be more involved in the team approach. The key is realization that the patients' needs may be complex and interdependent. Often, treatment is not offered in a vacuum and ideally is offered within a coordinated system, however tightly or loosely defined, because patients might not be able to negotiate through other aspects of life planning due to devastating change and loss or to the cognitive disruptions (if any). The specific approach and training of the therapist may dictate whether psychodynamic, cognitive–behavioral, stress management, cognitive remediation, or more counseling/educational-type approaches are used, or are useful, for the needs of patients.

Rarely are the challenges of identification more apparent than in working with persons who may be of similar age, background, and/or vocation and who are reasonably healthy until some catastrophic loss occurs. Psychologists might need to use the identification process in a manner that does not cloud patients' individuality and uniqueness of response, so that the therapists maintain the therapeutic stance needed to work effectively while retaining subjective understanding. The task of simultaneous subjectivity and objectivity is, of course, not unique to the practice of psychotherapy in a rehabilitation setting. In the work of psychological rehabilitation, however, there is a constant and humbling realization of one's own humanity and vulnerability. This sense of personal vulnerability may tug at the heartstrings. With time, experience, supervision, and self-awareness, this task is made easier for psychotherapists. Of course, the remarkable ability to adapt may be a humbling experience for clinicians working with patients rallying against all odds to cope with tasks of daily life that have become monumental. The inspiration derived can fuel professional growth, enhance therapeutic work, and enrich personal meaningfulness of the many values that form part of therapists' identity.

See Also the Following Articles

Coping ■ Stress

Further Reading

Americans with Disabilities Act (1990). 42 USC 12101nt.

Christensen, A-L., & Uzzell, B. (Eds.). (1999). *International handbook of neuropsychological rehabilitation*. New York: Kluwer Academic/Plenum.

Eisenberg, M. G., Glueckauf, R. L., & Zaretsky, H. H. (Eds.). (1999). *Medical aspects of disability: A handbook for the rehabilitation professional* (2nd ed.). New York: Springer.

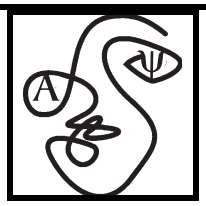
Frank, R. G., & Elliott, T. R. (Eds.). (2000). *Handbook of rehabilitation psychology*. Washington, DC: American Psychological Association.

Langer, K. G., Laatsch, L., & Lewis, L. (Eds.). (1999). *Psychotherapeutic interventions for adults with brain injury or stroke: A clinician's treatment resource*. Madison, CT: Psychosocial Press.

Lezak, M. D. (1995). *Neuropsychological assessment* (3rd ed.). New York: Oxford University Press.

Marinelli, R. P., & Dell Orto, A. E. (Eds.). (1999). *The psychological and social impact of disability* (4th ed.). New York: Springer.

Rosenthal, M., Griffith, E. R., Kreutzer, J. S., & Pentland, B. (Eds.) (1999). *Rehabilitation of the adult and child with traumatic brain injury* (3rd ed.). San Diego, CA: Academic Press.



Psychological Skills Training in Sport

Michael J. Mahoney and Benjamin P. Chapman

University of North Texas, Denton, Texas, USA

1. Introduction
2. Cognitive Skills
3. Emotion Self-Regulation Skills
4. Motivational Skills
5. Social and Interpersonal Skills
6. Training Psychological Skills
7. Conclusions

Further Reading

GLOSSARY

anxiety An experience characterized by fear, doubt, or worry; at a physical level, it is often felt as a racing heart, a nervous stomach, jitteriness, or jumpiness.

arousal A physiological state characterized by heightened activity of the sympathetic division of the autonomic nervous system; it prepares an organism for action and is a common, if not necessary, component of most athletic experiences.

efficacy expectations The belief that one is capable of performing a given behavior in such a way as to bring about a desired outcome; experienced athletes may hold higher efficacy expectations about their abilities to perform various skills and tasks in sport.

flow A state of consciousness in which one's attention and concentration are focused on an activity with precision and clarity; the performance of the activity is experienced as effortless and pleasurable and is often superior by objective standards.

individual zone of optimal functioning (IZOF) The intensity magnitude of a given biopsychosocial state in which an individual athlete performs best; IZOFs exist for the behavioral, cognitive, affective, motivational, bodily, operational, and communicative domains.

imagery The internal experience of sensory stimuli that are not externally present; for instance, visualization involves seeing something through "the mind's eye" and is a skill often used by athletes to calm anxiety, focus form, and facilitate performance.

mastery experiences Episodes of behavior in which a desired outcome is achieved through the actor's own effort; they lead to a sense of accomplishment, and when they accumulate, they facilitate a sense of competence and encourage efficacy expectations.

mental rehearsal A set of activities wherein an athlete imagines himself or herself performing some skill or task or behaving in a desired fashion; although this may involve imagery, elite athletes may mentally rehearse other components of an experience such as what it will feel, sound, and even smell like to perform an action.

psychological skills A set of techniques and strategies used to optimize athletic performance; they are largely nonphysical in nature.

self-instruction A process that encompasses internal reminders, reassurance, and self-talk; it may be used as a self-regulatory tool and is often effective in the remediation of some of the negative emotional states that athletes commonly experience.

©2004 Elsevier Inc.
All rights reserved.

self-monitoring The process of attending internally to one's ongoing cognitive and emotional processes as well as externally to one's behavior and performance.

self-regulation A process that is dependent on self-monitoring and involves evaluating one's degree of success in achieving desired goals as well as implementing both internal and external feedback to make necessary corrections; self-monitoring and self-regulation in sport performance appear to be skills associated with athletic success.

Psychological skills in sport constitute a set of techniques and strategies used to optimize athletic performance. Psychological skills have traditionally been neglected in favor of physical skills in assessing athletic performance. However, increasing evidence suggests that psychological skills may be just as important as their physical counterparts, particularly in elite-level athletes. Psychological skills in sport comprise different practices and strategies across the realms of behavior, cognition, emotion, motivation, and interpersonal functioning. Some techniques may facilitate the enhancement or development of various skills in athletes; however, much of what is known about psychological skills has been gleaned from studies of elite athletes, and research on skills training is not yet extensive. Therefore, caution is recommended in attempting to train psychological skills in a "one size fits all" fashion. However, interventions may be more confidently "targeted" to a specific group or, better yet, can be "tailored" to a specific individual.

1. INTRODUCTION

Andrea shakes out the tightness in her arms and shoulders while she waits in the shadows outside the gym's side entrance. She peeks around the corner and is hit by a gusty winter breeze from the open doors across the floor. The murmuring crowd looks like a tidal wave, rising high up into the rafters, waiting to crash down on her. She has competed in the championships before, but never in front of this many people.

Suddenly, the loudspeaker crackles, and the master of ceremonies announces a 5-minute break to check equipment before Andrea is due to compete. Her heart surges, and suddenly her limbs feel stiff and her skin feels clammy. She had been all ready, psyched up, and in the perfect rhythm. Now, she will get cold and go out there and make a fool of herself in her biggest meet ever.

Andrea thinks of her coach. He is not there, but she knows what he would say. She takes a deep breath to slow down her heart, remembering how well she has done in big meets before. Her practices have been going better than ever lately. Didn't she tell her friend just last night that her confidence was at an all-time high? Andrea closes her eyes and imagines what she will be doing in 5 minutes, what her body will feel like as it goes through the motions that are by now second nature to her. Her heart begins to slow, but at the same time, she bounces on the balls of her feet, does pushups off the wall, shakes her legs out. Blood returns to her limbs, replacing the clammy, paralyzed feeling. She takes another deep breath.

Andrea's name buzzes over the loudspeaker, but it seems distant, just loud enough to reach her. She walks into the gym, concentrating only on the equipment and imagining what her performance will feel like. The crowd begins to disappear; she hears only the rhythm of her own breathing and sees only the equipment. With a calm mind, she steps into the competition area, her body feeling like a compressed spring that is warm and ready for action.

In a minute, it is all over. Andrea steps back and looks up. The judges nod approvals, but she already knows that her performance was a personal best. The crowd's cheers spill down on her now like fresh air, no longer a suffocating tidal wave. She thinks of her coach and smiles.

This scenario is one example of many performance situations that confront athletes. Naturally, the situational demands of athletics vary largely depending on the sport. Is it an "open" sport, where the challenges and demands are constantly changing, or is it a "closed" sport, where the tasks are performed against a relatively stable backdrop (e.g., the green on that troublesome fourth hole does not suddenly change its size or slope, despite the frustrated impressions of golfers)? Is the athletic endeavor an individual one, such as wrestling or weightlifting, or a team game, such as soccer or basketball?

Research on elite athletes has always studied the psychological dimensions of athletic performance. What makes an athlete a good athlete? Given equal physical capabilities, what makes one athlete better than another? What makes a good athlete an extraordinary one? Contemporary sport psychology emphasizes a variety of factors that may influence performance differently across individuals and situations. For a

time, however, the presumably stable personality traits of athletes were thought to play the largest role in their success at sport.

1.1. The Athletic Personality

During the 1960s and 1970s, research examined personality differences between athletes and nonathletes as well as between successful athletes and unsuccessful athletes. Was there a general "athletic personality"? Were athletes more extraverted than their nonathletic peers? Were elite athletes a "different breed" from their less successful counterparts—more disciplined and aggressive? To explain the personality type associated with good athletes, Morgan proposed a "mental health" model suggesting that athletic success is inversely related to psychopathology. Other evidence gathered through the use of conventional personality inventories suggested that successful athletes tended to be more assertive, dominant, aggressive, self-sufficient, and vigorous; to have a higher need for achievement; and to be less anxious than less successful athletes.

Although such personality factors seemed to explain some of the differences between more successful athletes and less successful athletes, research of this sort fell out of favor by the mid 1970s. Sport psychologists grew more interested in differences between individuals rather than group commonalities. Successful athletes vary widely in characteristics such as pain tolerance, optimal arousal levels, and anxiety, and these factors interact in different ways across different athletes in different situations. So, if there is no general "athletic personality" solely responsible for one's performance in sport, what accounts for athletic success?

1.2. Psychological Skills in Sport

The answer appears to lie in the broad array of psychological skills reported by successful athletes. Surveys of elite athletes and studies of peak performance reveal a variety of mental, emotional, and motivational techniques that appear to account for a large part of success in sport. Evidence indicates that elite athletes employ mental imagery and bodily feedback, generate and maintain motivation, handle performance anxiety, concentrate, and retain self-confidence more effectively than do their less successful counterparts. Furthermore, such skills may be at least partly modifiable.

This article considers psychological skills in five broad categories: those skills that are primarily bodily and behavioral, those that are cognitive, those that involve primarily emotions, those that concern motivational states, and those that encompass interpersonal functions in sport. Bodily and behavioral skills relate to controlling one's physiological functions and actions such as regulating one's bodily state and readiness and acquiring and polishing motoric routines that are necessary for successful performance. Cognitive skills encompass a broad class of activities such as using mental imagery, monitoring one's thoughts and cultivating effective self-talk, and attending and concentrating in optimal ways. Emotional skills typically involve the regulation and optimization of affective states, including the ability to control anxiety and foster positive mood and feelings about oneself. Motivational skills provide athletes with inducement or impetus to excel and may involve things such as goal setting and responding to social facilitation. Finally, interpersonal skills relate to one's smooth interaction with others in sport and include communication abilities and the regulation of emotion in a team environment. Many skills function across several of these domains. The current classification scheme is intended only for ease of presentation and description and not as a discrete taxonomy.

1.3. Bodily and Behavioral Control Skills

Behavioral and bodily control skills operate at the physical level to enhance athletic success. Two primary skills involve regulating one's level of arousal and practicing or physically rehearsing motor skills.

1.3.1. Physiological Arousal Regulation

Physiological arousal refers to a state of heightened activity within the autonomic nervous system. High levels of arousal bring on the so-called "fight or flight" response, a reaction to threat and stress inherited from humans' distant mammalian forebears. Athletes may experience physiological arousal through the distinctive "charged-up" feel of adrenaline, increased flow of blood to skeletal muscles, increases in heart rate and respiration, and so forth.

It is important to note that such physical sensations often accompany the emotional experience of panic or anxiety. Although anxiety usually involves a strong component of physiological arousal, the two phenomena are not identical. An athlete may be in a state

of high physiological arousal but not be anxious; for instance, he or she may be angry. Similarly, an anxious athlete may experience low levels of physiological arousal, particularly if he or she is exhausted but still worried.

Some have suggested that, at an optimal level, arousal may actually diminish anxiety and/or serve to focus selective attention. For these reasons, different athletes often prefer different levels of arousal. One useful way in which to understand this relationship is from the perspective of an individual zone of optimal functioning (IZOF), an idea proposed by Hanin.

An IZOF corresponds to the intensity magnitude of a biopsychosocial state within which a given athlete feels most capable of executing the skills that are necessary for success in his or her sport. Examining autonomic arousal from this perspective, athletes must discern the levels of arousal at which they perform best. This may vary across behaviors, skills, and situations. Athletes who are aware of their somatic IZOF are able to detect when they are over- or underaroused and then can take corrective measures to regulate their arousal.

The ability to self-regulate arousal levels is critical for athletic success. This may take the form of last-minute mental or physical rituals to increase or decrease arousal, as discussed previously, or may be a more gradual process drawn out for hours (or even days) prior to a competition. In fact, the regulation of arousal can be important even during practice sessions. The ability to "turn on" or "turn off" the autonomic nervous response is largely a learned skill that is discussed later in the article.

Finally, different sports require different levels of physiological arousal. Games that are dependent on fine motor skills—archery, golf, shuffleboard, billiards, and so forth—clearly require a relative degree of calmness for effective skill execution. Sports that require strength, gross or explosive motor movements, and combative maneuvers—football, wrestling, sprinting, boxing, and so forth—generally necessitate higher levels of physiological arousal, although again optimal zones vary by individual athlete. It is also possible to be overaroused even in such sports; for instance, a soccer player may lose what is called "touch" in his or her passes and shots, instead kicking the ball into the stands.

1.3.2. Behavioral Practice

Motor learning of sport-specific skills requires physical repetition. Different degrees and types of repetition increase different skills at different rates in different athletes. But whether skill acquisition is sudden or gradual, the adherence to some sort of practice regimen is nearly always necessary for the learning, maintenance, and continued development of athletic skill.

One consideration is the amount of practice. Early research on the amount of practice in motor performance suggested that plateaus or ceilings sometimes may be reached, but modern work on deliberate (i.e., effortful) practice suggests that sustained practice generally leads to continued improvement, even over long periods of time. The relationship between practice and skill acquisition may or may not be linear and is usually a function of the athlete, the skill, and the degree of learning that has already taken place.

The variability or breadth of practice is also an important consideration. To learn a complete repertoire of sport-specific motor skills, an individual must practice the complement of skills that the sport requires. Sometimes, it is also beneficial to practice under a variety of conditions, particularly if the conditions of competition vary (as in changing weather patterns or variable lighting). Thus, varying the motor behaviors themselves and the conditions under which they are executed may lead to greater degrees of overall skill acquisition.

Finally, performance on complex motor tasks may be improved by decomposing them into simpler component movements. An athlete practicing a partial movement may focus attention and effort more thoroughly. Thus, proficiency acquired at these task parts may then translate into greater proficiency at the whole task. Part task practice typically works best for behaviors in which component parts are not highly interdependent; complex behaviors in which components are more highly organized and coordinated may benefit less from decomposition into separated constituents.

2. COGNITIVE SKILLS

Cognitive skills are those that primarily involve mental processing. This section discusses three categories: attention and concentration, mental rehearsal and imagery, and self-monitoring.

2.1. Attention and Concentration

Within the realm of sport, attention and concentration are important in many ways. Athletes need selective

attention to extract relevant features from the field of activity. Likewise, divided attention may be called on to follow diverse streams of action, particularly in open sports. In the earlier example, Andrea must monitor processes inside herself, such as her bodily feelings and sensations, as well as the events in the competition area. Without good divided attention skills, she might have missed her name when it was called over the loudspeaker.

Attention and concentration skills involve focusing mental effort on task-relevant cues and filtering out distractors. Some evidence has linked attention to optimal levels of arousal, and it has been suggested that the superior attention skills of elite athletes are related to their ability to regulate their own arousal levels. Optimal levels of arousal may narrow or focus attention on task-relevant details in the environment. Conversely, under- or overaroused athletes may find themselves to be distractible. But attention may also be focused too narrowly, with less successful athletes overfocusing on minute details of competition and not seeing "the big picture." In addition, less successful athletes may tend to allow their attention to focus on negative perceptions of themselves or their performance.

In this vein, Moran has classified challenges to attention and concentration as arising from either external distractors or internal distractors. External distractors arise from the environment and generally involve things such as crowd noise, equipment and playing conditions, verbalizations and strategic interference from opponents, and visual distractions. Internal distractors involve things such as misplaced or excessive thoughts about the past or future, waning motivation and fatigue, and anxiety. Attention and concentration skills permit athletes to screen these distractors and focus on relevant cues and situations developing during a practice session or competition.

Also relevant is Nideffer's model of attention, which proposes that attention may be either internally directed or externally directed and be either broad or narrow in scope. Different athletes may have preferred attention styles, and different sports may have different attentional demands. For instance, in soccer, an individual must direct attention largely externally and in a broad fashion to teammates, opponents, and other outside task demands. Optimal attention is a function of how well the individual's preferred attention mode fits that demanded by the sport. From a psychological skills perspective, successful attention during sport may be contingent on an athlete's ability to detect and implement the sort of attention required.

Some research has also found that elite athletes tend to use associative, as opposed to dissociative, attentional strategies. Associative strategies involve an active comparing of task demands with one's physical capacities—a continual comparison of external and internal perceptions. Dissociative strategies, on the other hand, involve separating external task performance from kinesthetic and somatic feedback, that is, "shutting out the pain" or distracting oneself from unpleasant bodily sensations. For instance, the fact that elite distance runners report using associative attention (whereas mediocre runners report using dissociative strategies) suggests that superior performance may in part be a function of matching athletes' physical abilities to external task demands—a process that is difficult, if not impossible, if athletes have "tuned out" their bodily signals.

The ultimate form of concentration occurs during episodes of peak performance labeled as flow. Peak performances are episodes in which athletes report complete engagement with their immediate tasks, sometimes yielding results of an incredible nature. Several features of peak performance episodes appear to be noteworthy. Athletes in flow report that all things irrelevant to the immediate task tend to recede from awareness. The spectators, distracting thoughts, and indeed even pain may disappear from athletes' perceptions. All available resources are focused unswervingly on the athletic performance. Episodes of peak performance are characterized by such supreme concentration that even an awareness of time may be diminished or lost. Finally, athletes reporting flow experiences often say that they feel as though their performance is executed with an automated ease.

2.2. Mental Rehearsal and Imagery

Another skill that clearly contributes to athletic success is mental rehearsal. Mental rehearsal refers to the practiced imaging of successful athletic performance. Imagery refers to the specific thoughts, sounds, smells, visual array, and sensations that an individual experiences when he or she conjures up a mental experience of something. Mental rehearsal and imagery, particularly of specific skills, may facilitate motor learning. Repeatedly imagining successful athletic execution may enhance an individual's certainty that he or she is capable of performing a given skill. In contrast, evidence suggests that although successful imagery is related to better performance, imagining unsuccessful outcomes tends to be a detriment to performance. Elite athletes

also report that the images they employ are extraordinarily clear, vivid, and multisensory—nearly lifelike.

Imagery is not confined to visualization. Many elite athletes report using a holistic form of imagery in which they imagine the sounds, feels, and even smells of a situation. Elite athletes also report using "first-person" imagery, that is, imagining successful outcomes from their own perspectives. Less skilled athletes may be more apt to use external imagery in which they view themselves performing some action from the outside as if they are watching a video of themselves.

2.3. Self-Monitoring

The execution of physical skills is only the beginning of athletic performance. Success in sport also involves monitoring one's performance, evaluating it accurately against expectations and task demands, and then changing or maintaining it accordingly. Also required is the awareness and regulation of thoughts, feelings, and even physiological arousal. Self-monitoring is the ability to accurately perceive and evaluate performance (or internal experience) against task demands, whereas self-regulation is the ability to manipulate one's behavior (or internal states) accordingly.

Some evidence suggests that elite athletes are better able to self-monitor and self-regulate than are their less successful peers. In evaluating performance against demands, individuals with high mastery of the tasks at hand (i.e., more successful athletes) seem to notice and use feedback from both positive and negative aspects of performance. In contrast, individuals who are low in task mastery (i.e., less successful athletes) seem to benefit only from focusing on the successful dimension of their performance. This does not imply that to become a better athlete, one must immediately begin focusing on performance failures in addition to successes. On the contrary, such a focus may be counterproductive to self-confidence and inappropriate in less advanced athletes.

In addition to evaluating one's behavioral outcomes against those called for in a sporting situation, self-monitoring often involves tracking thoughts. Negative thoughts often exert deleterious effects on performance. In the earlier example, Andrea experiences doubt about her ability to deal with the unexpected delay in the form of a sudden thought. The awareness of such thoughts, called negative self-statements, is extremely important in athletic performance, particularly because many people are not immediately conscious of them.

3. EMOTION SELF-REGULATION SKILLS

This section focuses on two areas that are closely related to functions of emotion: the relationship of anxiety to athletic performance and the nature of self-confidence and self-efficacy.

3.1. Anxiety and Personal Meaning

For many years, performance anxiety in sport was considered to be debilitating, that is, to be avoided at all costs. More recent perspectives suggest that such blanket conclusions are somewhat oversimplified. The nature and effects of anxiety vary highly across individuals and situations. To be sure, performance anxiety often hampers otherwise talented athletes, and some athletes manage it better than do others. The worried emotions that are characteristic of anxiety often have cognitive and physiological counterparts that may be controllable through self-talk or relaxation. But perhaps most important in the management of anxiety is the meaning that athletes attach to it.

For some athletes, anxiety represents a fundamental fear of failure that has deep personal ramifications. Many athletes attach their identity or self-worth to their athletic abilities and performance. In this case, failure in sport is tantamount to failure as persons, a shortcoming that is much more profound and fundamental than merely losing a game or performing poorly during a competition. Poor performance, from this perspective, is a destructive event to be feared and guarded against at all costs, yet its mere potential is so threatening that athletes can never ignore or forget it. In instances such as this, anxiety is debilitating precisely due to the meanings attached to it.

Although most athletes report that high levels of anxiety are not pleasant, anxiety may be facilitative as well. Research with some elite athletes suggests that they view their anxiety as an ally or a companion. For them, insufficient anxiety may be disconcerting because it signals a lack of interest or engagement in the task at hand. For these athletes, management of anxiety involves "pacing"—allowing a slow buildup of anxiety, beginning days or even weeks before a competition, that spurs them to maximal performance. Rather than eliminating anxiety entirely, athletes may be aided by developing an understanding of the function that anxiety plays for them. Once this has been achieved, it will be easier to know how to regulate anxiety so that it

may be kept at levels that are facilitative, but not inhibitory, of performance.

3.2. Self-Confidence and Self-Efficacy

Self-confidence is a general feeling of competence or certainty in one's ability to succeed. Self-efficacy refers to the belief that one can successfully perform a specific action. Both are critical to athletic performance. Because self-confidence is so vague and general, the descriptions in this section note the components that recent research suggests constitute self-confidence in sport as well as a general framework for understanding the role of self-confidence in athletes.

Vealey proposed a model in which self-confidence in sport may be seen as a product of nine dimensions: mastery, demonstration of ability, physical/mental preparation, physical self-presentation, social support, coaches' leadership, vicarious experience, environmental comfort, and situational favorableness. These nine areas can be grouped into three domains: Achievement (involving mastery of sport skills and the demonstration of an athlete's ability), Self-Regulation (encompassing an athlete's degree of physical/mental preparation and feelings about physical self-presentation), and Climate (including an athlete's sense of social support, perceptions of coaches' leadership, favorable vicarious experiences, and comfort with the environment and degree of situational favorableness). To the extent that these factors arise or are arranged in a more advantageous way, an athlete will feel more confident in sport.

At a broader level, self-confidence is not merely the absence of doubt or the rational certainty of success. Like anxiety, self-confidence may be deeply tied to athletes' personal meaning systems. Relative to their less skilled peers, elite athletes may experience a more flexible connection between their self-worth and their sport performance. The result may be a decrease in felt performance pressure because their value as persons is not totally invested in any particular performance. In other words, elite athletes may have more ability to remain "faithful to themselves" (as suggest by *confidere,* the Latin root of "confidence") regardless of the outcome of their sport performance.

Successful athletes also generally hold high efficacy expectations. Efficacy expectations are anticipations of achieving desired outcomes or goal states. Efficacy expectations may be increased through the praise and assurance of others as well as through techniques such as mental rehearsal and imagery. But efficacy expectations are usually developed through repeatedly performing a graduated series of tasks more and more successfully. This process is called mastery.

When an individual performs some action successfully, it serves as a mastery experience. A mastery experience involves executing a behavior that leads to a desired goal state, with the result being an increase in the athlete's confidence that he or she can perform the action successfully in the future. Practice in sport allows athletes to build self-efficacy through potential mastery experiences—but in a safe environment where failure does not result in the loss of competitions. Successful performances in prior competitions may be even better mastery experiences because they involve achieving goal states under conditions similar to future competitions. In the earlier example, although Andrea has never performed before such a large crowd, she has competed previously under conditions that were similar in other ways.

4. MOTIVATIONAL SKILLS

This section discusses two types of motivation common in athletics: goal setting and social facilitation. Goal setting usually serves long-term purposes, whereas social facilitation is generally an external motivator on which some athletes capitalize.

4.1. Goal Setting

Goal setting marshals and directs an athlete's energies and efforts either in the short term, as in Andrea's intention to win her immediate competition, or in the longer term, as in her plans to increase her conditioning during the off-season. Successful athletes generally report using personally meaningful goals. Choosing goals with little personal relevance is likely to produce diminished effort. In pursuing their goals, elite athletes also tend to report a healthy balance between their athletic pursuits and other dimensions of their lives. Athletes who drive themselves single-mindedly may generate momentum in the short term but often tend to ultimately lose balance in their lives. Because of this, they may be more susceptible to slumps in motivation and performance.

In general, research suggests that performance benefits more from specific and/or realistic goals than from vague and/or unrealistic goals. Likewise, some evidence suggests that intrinsically motivated athletes

may improve when allowed to set their own goals, whereas extrinsically motivated athletes may improve when others set their goals. Another area of research has examined the effectiveness of goals aimed at task mastery versus goals aimed more externally at competitive superiority. Individuals motivated by task mastery challenge themselves in ways that develop their skills, whereas individuals motivated primarily by outcomes may choose goals or activities that are easily accomplished to avoid threats to self-esteem. It is unclear whether task mastery goals lead to better performance than do outcome-oriented goals, but athletes who employ task mastery goals do seem to experience greater increases in self-efficacy than do those who use outcome-oriented goals.

4.2. Social Facilitation

Another motivational source in athletics lies in social facilitation. Social facilitation is a process whereby performance is increased due to the presence of others. It is one of the oldest areas of sport psychological research. Better performance in the presence of competitors or spectators has been explained by classical psychological concepts such as drive theory. Spectators tend to increase arousal, and drive theory proposes that arousal increases the performance of dominant responses. This is a desirable state of affairs when dominant responses are correct and well learned, as in familiar tasks. But the presence of spectators may hamper performance when dominant responses are incorrect or underlearned, as in novel or unfamiliar tasks. Therefore, the presence of others may or may not facilitate performance. Not surprisingly, research on drive theory and social facilitation in sport has produced some inconsistent results. Social facilitation effects appear to vary across real-life sports, where motor behaviors are more complex than those studied in laboratories.

Interviews with elite athletes suggest that the ability to draw inspiration from a crowd or the presence of opponents may be highly individualized. For some, the attention may be energizing. Such individuals generally benefit from some level of awareness that others are present and watching. In this case, the crowd may become a task-relevant cue in competition, and athletes may develop the ability to capitalize on social facilitation by attending to the crowd. Performing in front of individuals in practice may help to develop the skill of drawing energy from a crowd as well as provide practice in partitioning attention among motor performance, internal cues, and the external importance of the crowd. For other athletes, knowledge that they are the center of attention becomes anxiety provoking. Such athletes may benefit more from skills training that minimizes the scrutiny felt from the presence of others.

5. SOCIAL AND INTERPERSONAL SKILLS

A broad array of social and interpersonal skills may be brought to bear in athletics. Smooth and effective interactions with others—not just teammates but also coaches, officials, and even opponents—may facilitate higher degrees of performance. This section considers two areas that are relevant to interpersonal skills: communication and interpersonal and intragroup anxiety in a team context.

5.1. Communication

Being able to communicate in sport is often critical to success. This is true even in the case of individual sports, where an individual must still interact with officials, coaches, and opponents. In general, performance will be maximized to the extent that an individual can effectively encode and deliver feedback and messages to teammates, coaches, and officials as well as decode and incorporate received information.

In team sports, delivering encouragement and technical feedback to teammates is often essential. Sending understandable messages at appropriate times permits teammates to correct deficiencies or other performance impediments, facilitating better team play. Being able to communicate openly and effectively with coaches during training and practice also permits the ready flow and interchange of necessary (and often vital) information about an athlete's functioning across biopsychosocial states. Finally, communicating appropriately with officials during the heat of competition is a particularly difficult, but nonetheless indispensable, skill.

Communication also involves receiving messages. Often, feedback may be delivered from teammates, coaches, opponents, or officials in an emotional or insulting manner. The ability to evaluate feedback for its truthfulness and usefulness, and then to glean helpful elements while rejecting unhelpful parts, is often critical to improving sport performance. Athletes who "can't take criticism," or who incorporate all critiques

of their performance without discrimination, may be missing out on valuable opportunities to benefit from appropriate guidance.

In addition to articulating thoughts clearly, communication patterns also affect the emotional climate. Belittling a sensitive or anger-prone teammate might not be an effective performance enhancer for the team, and communicating with a high level of tension in an already pressurized competition might serve only to increase the anxiety of some teammates. In some instances, the contagious effect of high-level emotions may have an overall positive impact on team play. In other cases, it may be disastrous, as in the case of raising individuals' anxiety levels.

5.2. Interpersonal and Intragroup Anxiety

The uniquely social nature of sport may serve to heighten anxiety. Although individual traits or dispositions contribute to anxiety, Hanin described two types of anxiety that are also partially created by the team or interpersonal environment: interpersonal and intragroup anxiety. The two are related and, whether operating alone or in conjunction, may exert deleterious effects on individual and team performance.

In interacting with a training partner, a coach, or some other individual, interpersonal anxiety may arise. Interpersonal anxiety is typically a state in which an athlete fears negative evaluation from, is distrustful of, or is uncomfortable with the partner of an interaction. This may occur commonly with an unfamiliar or new coach or in communicating with a higher status team member. There is some indication that an athlete may feel less or no interpersonal anxiety in interacting with a cohort of equal or lower status because social comparison works favorably for the athlete in such cases.

A related form of anxiety, intragroup anxiety, may also exist in team contexts. In intragroup anxiety, the broader social milieu of the team itself affects an athlete's subjective anxiety. Intragroup anxiety may arise as a result of things such as rejection or poor treatment by teammates, a lack of group cohesiveness or poor team atmosphere, a lack of experience, and poor self-confidence. Athletes who are naturally more anxious possess greater degrees of negative emotionality and are more concerned with maintaining socially desirable appearances and behaviors. They may be more prone to intragroup anxiety. Lower status group members may be particularly susceptible to this type of anxiety.

The anxiety management skills discussed in the next section apply to interpersonal and intragroup anxiety. But in addition, trainers, teammates, and coaches may play important roles in regulating the interpersonal and intragroup anxiety of individual athletes. To the extent that familiarity and comfort can be enhanced between the members of a dyad, interpersonal anxiety may decrease and performance may increase. In a broader team setting, a mature group climate and healthy group dynamics may be encouraged by perceptive and effective coaches.

6. TRAINING PSYCHOLOGICAL SKILLS

Some evidence suggests that arousal regulation, behavioral practice strategies, attention and concentration, mental rehearsal, self-instruction, appropriate goal setting, and communication skills can be trained. However, in contrast to conventional psychotherapy where accountability has long been an issue, research on psychological skills training in sport is still in its infancy. Much of what is known about psychological skills has also been gleaned from studies on elite athletes, and the techniques and strategies observed in elite athletes may or may not be effective in less skilled athletes. Thus, the consequences of attempting to teach a general canon of psychological skills in a "one size fits all" approach may range from success for some athletes, to wasted effort for other athletes, to perhaps even detrimental effects for some individuals. Therefore, overzealous or indiscriminate training of all skills with all individuals is discouraged in favor of interventions that take into account an individual athlete's personal characteristics, needs, sport, and levels of physical, cognitive, and emotional development. With this caveat in mind, this section reviews some commonly used protocols and techniques for developing psychological skills.

6.1. Arousal and Anxiety Regulation

Because a large component of anxiety is physiological overarousal, the following considerations are relevant to the training of both anxiety and physiological arousal regulation. For anxious and overaroused individuals, various forms of relaxation training are the interventions of choice.

One form of relaxation training is called progressive muscle relaxation. Progressive muscle relaxation involves systematic decrements in the tension of the body's skeletal muscles. Typically, an individual flexes and then relaxes one muscle group at a time, often using visual imagery to facilitate the effect. Breathing becomes a focus, with special attention paid to a slow and steady rate of breath and deep inhalations and exhalations. When executed properly, these processes slow heart rate and decrease neuronal firing in the skeletal muscles, with parallel arousal decrements in the sympathetic division of the autonomic nervous system.

The gist of progressive muscle relaxation may be learned in a single session with a skilled facilitator, but mastery requires regular practice. A script or tape-recording may be employed to guide the progressive muscle relaxation, often beginning with arm and shoulder muscles and proceeding downward (although other variations are common). Employing tapes enables coaches, trainers, and athletes to institute relaxation practice without the presence of a trained facilitator, allowing relaxation sessions at any time or location and at less expense than bringing in a consultant. A variety of relaxation tapes are commercially available.

Another training mechanism for arousal regulation is biofeedback. Biofeedback involves electronic monitoring of an athlete's physiological signs, especially heart rate, pulse, galvanic skin response, and muscle tension. This information is subsequently used to guide progressive relaxation. With biofeedback data, problem areas in the relaxation process can be pinpointed and targeted with precision. After biofeedback training, an athlete often gains control over physiological processes that were previously beyond his or her ability to regulate. This permits an athlete to relax with greater ease and efficiency.

Biofeedback training generally occurs over multiple sessions. Initially, an athlete is introduced to the equipment and basic principles of biofeedback. He or she is shown, in visual terms, the effect of small bodily changes on heart rate and other indicators of arousal. Initially, an athlete may find it difficult to exert control over bodily processes that occur automatically. However, over successive practice sessions in a biofeedback laboratory, he or she learns how the manipulation of subtle and previously ignored physiological operations may affect levels of arousal and felt anxiety. Eventually, the biofeedback equipment is no longer needed, and an athlete can effect the necessary alterations to his or her arousal level in other contexts.

Meditation training may also reduce arousal and anxiety. Meditation involves relaxation techniques such as deep breathing but adds an element of mental focus, usually in the form of a thought or an image that is repeated or on which attention is focused. In this way, meditation may also serve to train mental imagery abilities and attention and concentration processes while also regulating physiological arousal. In learning meditation, an individual focuses on relaxing physically through deep breathing and gently releases competing thoughts, mental chatter, and worries while constantly returning to a "centering" image. This is best learned from either an experienced meditator or the use of tapes.

In developing meditation skills, regular practice is imperative, and patience and gentleness are critical, because it is extremely difficult to detach from "mental chatter." Finding a regular time and place for meditation practice often augments the process, but it is important that an athlete experience the setting for meditation as comfortable and peaceful. The duration of meditation practice is often specific to individual tolerance levels, but early practice sessions generally should be kept to shorter lengths (10–20 minutes for most beginners).

Other interventions permitting arousal and anxiety regulation may include prayer and hypnosis. Although one does not typically think of prayer as a sport performance skill, prayer may facilitate relaxation and focus attention. It may also provide an athlete with a calming reassurance that a divine force is operating on his or her behalf. Similarly, hypnotic states are defined by nearly total physical relaxation. However, not all individuals are easily hypnotizable, and trained hypnotists are not always available. Thus, unless an individual has access to a certified instructor, it is a less practical form of intervention and not an easily acquirable skill, although some individuals are able to learn to hypnotize themselves.

Finally, arousal regulation also involves "psyching up." Athletes may be trained to drastically increase arousal levels over a short time. This type of arousal regulation training has been less researched, but control may be exerted through physical channels (e.g., slapping oneself, shaking, jumping up and down) or through cognitive mechanisms (e.g., imagining anger-producing things, planning and anticipating challenges during competition, focusing on context cues such as the crowd, engaging in final rituals that have been repeatedly paired with high-arousal states in the past). Habitually underaroused athletes must explore

a variety of last-minute psyching strategies and must retain and practice those that seem to be most advantageous.

6.2. Part Practice

Behavioral practice skills and motor learning can be facilitated through part practice. In part practice, an athlete can focus greater attention and concentration resources on simpler components until they become automatized, thereby freeing up resources for execution of the whole task. This discussion considers three forms of part practice: segmentation, fractionation, and oversimplification.

In segmentation, a complex task is divided into a number of constituent dimensions, which are then trained separately. For instance, a gymnast training a routine on the balance beam may separately train the mount, actions performed on the beam, and the dismount. Often, repetition of task segments may involve graduated acquisition to more complex combinations of segments. In this way, task parts can be "chained" together, usually from the beginning pieces of a complex behavior to the end (forward chaining) or from the end pieces of a complex sequence to the beginning (backward chaining). The degree to which facility in task segments transfers to the whole task may be determined by whether the segmented practice involves integrated control of the "links" in the behavior chain. If an athlete does not learn necessary transitions between segments, the whole movement might not improve substantially even in the face of dramatic improvements in the segments.

In fractionation, an athlete trains separate pieces of a complex action that normally occur at the same time. Results regarding the effectiveness of this technique have been somewhat mixed, but it seems that for tasks in which component behaviors are highly integrated because they occur simultaneously, practicing various parts in isolation often fails to develop coordination among them. Research investigating fractionation does suggest that it may be useful for complex tasks where subcomponents must be randomly combined, that is, where there is not a single template or control/coordination process for the task. For instance, in a video game, multiple fractionated skills (e.g., making a figure run, jump, or fire a gun) may be called on, but in random combinations.

A final form of part practice is simplification. Simplification generally takes the form of training easy versions of complex motor behaviors as precursors and building blocks. A juggler may first learn to juggle two scarves, then progress to three baseballs, and then finally to five or more torches or knives. In this example, the degree of difficulty (as well as the danger of potential mistakes) increases incrementally. Again, research on simplification has been somewhat mixed. Although it seems to be an effective teaching technique, there may be a point at which simplification of a complex motor behavior alters the essential nature of a task. At this point, it is unlikely that skill acquired at the simplified task would generalize to the more complex task. However, if the fundamental control processes of the complex task are not altered to a great degree, simplification may facilitate motor skill acquisition in some athletes.

6.3. Improving Attention and Concentration

Attention and concentration skills are malleable and can be learned and improved through repeated practice. Several exercises have been proposed, including scanning number tables and staring intently at a sport-related object. However, the merit of such training techniques lacks systematic substantiation. Nonetheless, one technique that has gained some credence is adversity training.

In adversity training, the specific alterations to practice sessions are less important than their general effect: to create a training atmosphere in which distractors, unconventional circumstances, situations, and/or environmental interferences exist. By training in the presence of such potential disruptions, athletes develop a more robust ability to focus their attention and resources on task-relevant cues and screen out potential impediments that arise randomly in the competition environment. By treating practice sessions of this sort "like the real thing," athletes learn to screen out intrusive and novel stimuli as a matter of course so that if such nuisances appear in actual competition, they will be less disruptive. Although little systematic research has investigated adversity training, it is a manifestation of the well-established principle of state-dependent learning and anecdotal support is plentiful.

Another way in which divided attention skills may be developed is through dual-task practice. In sport, athletes are often required to perform multiple tasks simultaneously, and this requires switching skills to allocate attention between the different tasks.

Training multiple tasks simultaneously can help athletes to develop attention allocation strategies and increase attentional resources. For instance, a basketball player's dribbling skills may be improved by practicing dribbling two balls at once while looking ahead. In this way, dual-task practice facilitates automatization of component movements, thereby freeing an athlete's attention resources for allocation elsewhere.

However, it is important that dual-task practice be conducted with sport-specific tasks. Developing divided attention on an unrelated array of activities might not generalize to the idiosyncratic demands of a particular sport. Training attention skills through dual-task practice should also take into account an individual athlete's attention resources and skills. Individuals vary widely in their reservoir of attentional resources and in their skills at allocating them. But even for expert performers, continued overload of attentional resources through graduated dual-task or multitask practice may permit ongoing increases in attention skills.

Finally, a variety of other strategies have been proposed for improving attention and concentration skills. Preperformance routines, which generally consist of a series of motoric movements and/or accompanying thought and self-instruction, may serve to focus concentration. Centering exercises and mental imagery may also facilitate concentration on external task demands indirectly through augmenting internal focusing skills. Regulating arousal to optimal levels has salutary effects on attention or concentration. And although the ability to enter flow voluntarily seems to vary across individuals, certain strategies may help individuals to induce flow states. Setting miniature goals in practice sessions, such as executing a tennis serve with perfect form five times in a row, have been helpful at inducing flow in athletes. Transforming practice sessions into "games" may also help. For instance, soccer players practicing "juggling" the ball with their feet may elect to play a juggling game in which the player who can juggle the longest gets to skip a mandatory conditioning run.

6.4. Self-Monitoring and Self-Instruction Training

Self-monitoring may be facilitated through recording successful and unsuccessful behaviors in practice and competition. These records may then be reviewed either individually or with a teammate, coach, or trainer to examine patterns in an athlete's performance. With a log of successes and failures (and the assistance of another pair of eyes), athletes may begin to see performance patterns that can help them to pinpoint specific areas that need closer attention. Videotaping and reviewing performance may also facilitate self-monitoring skills. For instance, an athlete may realize that a particular type of performance mistake is always preceded by a motoric or kinesthetic flaw that can then be expunged during subsequent practice sessions.

Self-instruction may be taught through the use of verbal scripts. Scripts serve the purpose of providing a format, or template, of helpful statements and important reminders that an athlete can implement before, during, and after competitions. Self-instruction scripts ideally provide the athlete with a ready storehouse of important reminders and encouragers that, with practice, will become automatic. When well practiced, self-instruction may provide instantaneous correction, redirection, and reassurance in the midst of competition.

Cue words also may assist in self-instruction. Salient cue words, when implemented regularly, will serve to convey immediate corrective feedback to athletes when performance errors occur or are in danger of occurring. Cue words should be personally selected by an athlete but should be relevant and bear some important meaning about the activity being performed. For instance, a cricket bowler may employ the phrase "follow through" to increase the accuracy of his or her pitches. Common cue words may also be shared with teammates or coaches, who may assist athletes in acquiring and regularly implementing instructive words during performance.

6.5. Mental Imagery and Mental Rehearsal Training

Mental imagery is often conducted in states of relaxation, although it may be practiced anywhere. In general, however, the absence of distractions in relaxed or meditative states permits more efficacious use of mental imagery. Guided imagery is best learned with the help of another knowledgeable coach, trainer, or sport psychologist. Usually conducted during a state of relaxation, a series of relevant sensations are described in detail, invoking mental representations of sights, sounds, smells, and so forth. The images may proceed from an external perspective (i.e., a "camera reel" of a performer in action) or from an internal perspective (i.e., from the performer's view). Again, the content and use of imagery should be dictated by the individual athlete's needs. For an athlete who has trouble "getting

in the zone" prior to a competition, bodily sensations of tension and anticipation may be invoked during guided imagery practice, along with a "tour" of the precompetition environment and preparations, ending with an image of the athlete fully aroused, attentive, and eager to execute the actions that are necessary for competition.

Mental rehearsal of athletic performance may be conducted in nearly limitless ways. Some sport psychologists have developed structured training programs to help athletes improve mental rehearsal. Researchers have recently proposed a model that may be useful in the training and practice of imagery. Athletes are encouraged to consider where (e.g., during practice, before a competition) and why (e.g., trying to build general confidence, trying to improve a specific skill) they employ imagery. Athletes are trained not only to select places and times to use imagery that is consistent with what they are trying to achieve but also to use different types of imagery for general or specific goals. Imagery may be either specific (e.g., imagining the successful performance of a particular skill) or general (e.g., imagining oneself competing with confidence). The bulk of research suggests that imagery practice enhances specific skills and, like physical practice, should be distributed across time.

Visualization may be trained by using photos or videos as guides. Because real-life visual stimuli tend to be detailed and vivid, using them as models may facilitate visualization in beginners. Training may proceed from looking at sport-specific pictures or films, concentrating and noticing every detail, to engaging in mental practice wherein the model is recalled with as much vividness and detail as possible. If visual models of an athlete's own top performances are available, these may facilitate transition into personal visualization in practice. The content of visualization practice should match an athlete's goals. At a generic level, an athlete might visualize himself or herself preparing for, engaged in, and winning a competition. But at an individually specific level, an athlete might train visualization skills around a particularly difficult maneuver or motor sequence in performance. If motivation or confidence is an issue, visualization of goal states, successful outcomes, and mistake recovery may be helpful.

6.6. Acquiring Emotion Regulation Skills

Some techniques for reducing anxiety were addressed previously in the discussion of arousal and anxiety regulation. However, in addition to relaxation exercises, a variety of other strategies and interventions exist for regulating distressing emotions. These can often be developed using techniques found in most conventional psychotherapies. Although space constraints prohibit extensive coverage of these methods here, a variety of general self-help books provide thorough and accessible discussions of emotional regulatory strategies.

One way in which emotional distress may be remediated involves examining and challenging the thoughts associated with negative emotions. This requires that an athlete first develop an awareness of the sort of thinking that accompanies negative emotions. For instance, in the introductory example, Andrea experiences a surge of anxiety. Such a rush of nervous feelings may very well be accompanied by the thought, "If I don't perform well here, my career will be ruined." Naturally, this sort of absolutistic thinking serves only to heighten feelings of anxiety.

Identifying the self-defeating, irrational, catastrophic, and/or distorted thoughts behind negative emotions can often be difficult and frustrating for athletes learning emotional regulation because such thoughts may be so automatic that they are hard to detect. The use of logs, in which athletes track emotionally distressing moments and the thoughts beneath them, often helps to illustrate the connection between thinking and emotion and illuminates patterns. When athletes develop such awareness, they know where to intervene when emotions turn problematic.

Once the thoughts driving negative emotional responses have been reliably identified, the next step in emotional regulation is to begin systematically challenging such thoughts. Athletes can learn to "act as their own defense attorneys" in response to self-derogating thoughts or self-anger and to question the likelihood of fictional finalisms or the validity of hidden assumptions. However, this process is often effortful and requires extensive practice because negativistic thinking patterns tend to be quite entrenched. It is important to note that identifying and challenging self-defeating thoughts is likely to be an unfamiliar experience to an athlete learning emotional regulation. For this reason, the athlete might wish to enlist the aid of a personal therapist, counselor, or trained sport psychologist.

Self-confidence may also be trained through techniques such as imagery and mental rehearsal of confident performance, monitoring and replacement of self-defeating thoughts, seeking out mastery experiences,

and validation and encouragement from others. However, self-confidence in athletics may also be rooted to deeper feelings of personal inadequacy. When this is the case, personal counseling or psychotherapy is likely to be useful. In fact, for athletes experiencing emotion regulation problems at a deeper level and/or extending outside of competition, personal counseling or psychotherapy is highly recommended not just to improve sport performance but also to improve overall happiness and quality of life in general.

6.7. Training Appropriate Goal Setting

The consequences of inappropriate goal setting can be more deleterious to an athlete's motivation than training bereft of goals because the wrong type or set of goals can set up an athlete for failure. Research suggests various parameters that are most conducive to goal attainment, although goal setting should also be considered from an individualized perspective due to the wide variation in individual abilities.

In general, athletes should be encouraged to select a series of easy proximal goals early on (e.g., accomplishing small things that are close and immediate). Easy proximal goals offer the possibility of quick feedback and mastery experiences, which serve to enhance motivation. The arrangement of proximal goals may also form a chain toward some ultimate goal. For instance, a team may institute night-by-night goals during the regular season to reach the league championship. However, it is important to avoid an overly rapid or unrealistic series of proximal goals. Individuals or teams may be prone to set less realistic goals after successful achievement of early easy goals.

Athletes learning to set goals should also be encouraged to select specific goals. Specific goals permit achievement in precise areas relevant to success. For instance, rather than setting a goal to improve one's overall strength, a weightlifter may set a specific goal of improving his or her maximum clean and jerk by 2 kilograms and his or her maximal snatch by 1 kilogram. If the weightlifter also invokes the principle of proximal goal setting, he or she may resolve to achieve this at the end of 4 weeks of training sessions. Specific goals render large tasks (e.g., winning a competition) less imposing or intimidating by breaking down a desired end state into components.

Goal-setting skills are best developed by encouraging athletes to use mastery goals as well. Athletes should be encouraged at first to "compete against themselves" by improving past performance or meeting specific competence levels independent of competition. Developing athletes who set competitive goals may rob themselves of a sense of personal control in pursuit of their desired ends because competitive goals tend to be at least partly contingent on other individuals. The achievement of mastery goals, in contrast, is generally more within the direct control of the athlete himself or herself. Mastery goals, in turn, promote the development of self-efficacy more readily than do competitive goals.

An athlete may also wish to begin with individually chosen goals. In cases where the athlete is uncertain of his or her abilities or appropriate first goals, it may be helpful to develop goals initially with a coach, trainer, or sport psychologist in a collaborative fashion. In such instances, goals that are personally meaningful to a given athlete may serve as the foundation for early development. The mastery experiences attached to maintaining individually salient goals may be more motivating than the sense of accomplishment felt at meeting institutional or external goals, although an athlete who is extrinsically oriented may also benefit from external goals.

Finally, it is important to note that these considerations apply primarily to the training of goal-setting skills in inexperienced athletes. Elite athletes may benefit just as much, or in some cases more, from distal and competitive goals. But the use of proximate, specific, mastery, and individual principles can be highly recommended to less advanced athletes wishing to develop goal-setting skills.

6.8. Enhancing Communication

Hanin recommended three general strategies for enhancing communication skills in a team environment. First, if communication problems are a function of the specific subset of individuals involved, such as having an aggressive or belligerent athlete in the starting lineup, one may alter the composition of the unit. Communication patterns among players are likely to change accordingly. However, this is only a temporary solution and does not remediate the root of communication problems if it resides in one or more particular individuals.

If certain activities, tasks, or situations tend to produce communication difficulties, one may redefine, alter, or substitute tasks that do not evoke such difficulties. For instance, if one teammate always has trouble performing

in a certain drill and this leads to negative interactions with other players, an alternative drill that accomplishes the same thing may be instituted. Naturally in some situations, such as rehearsing a particular play that is indispensable to team performance, one cannot alter the situation or group task without compromising the integrity of practice. In such instances, communication problems may be dealt with by managing players' communication at a behavioral level.

At the level of communication behaviors, athletes may be taught to monitor and regulate the content, delivery, and even frequency of their messages. Through observation, instruction, and discussion, they may develop specific styles of communication for specific teammates. Negative feedback may be sandwiched between positive comments, accusing or derogatory content can be unlearned in favor of appropriate and task-relevant comments, and athletes can be taught active listening skills so that teammates feel that they are heard.

At both the individual and team levels, communication skills may be practiced through role-plays outside of practice. As new communication patterns are acquired, athletes should be encouraged to implement them during practice and to observe their own and others' reactions to these changes. The ability of an athlete or a team to transfer new communication skills to competition should then be monitored closely, with appropriate discussion of both positive and negative exchanges and necessary adjustments.

Whether communication deficits are a function of skill deficits or situations, facilitating communication skills in athletes requires close observation of problematic communication patterns, including their antecedents and effects. With such knowledge, coaches and/or other helpers may meet with individual team members, as well as the team itself, to discuss specific problems. It is also important to provide feedback on instances of effective communication as well so that appropriate interaction patterns can be reinforced.

7. CONCLUSIONS

Research on elite athletes reveals that rather than a general "athletic personality," psychological skills are more important determiners of athletic success. These skills are mostly acquired rather than innate, and they span the bodily/behavioral, cognitive, emotional, motivational, and interpersonal domains. Bodily/Behavioral control skills encompass arousal regulation and behavioral practice of motor skills. Within the cognitive realm, mental rehearsal, attention and concentration, and self-monitoring and self-instruction skills are critical to athletic success. Emotional balance, controlling anxiety, and maintaining self-confidence are necessary. Motivational skills involve setting appropriate goals and capitalizing on social facilitation. Finally, interpersonal skills involve effective communication patterns and management of emotions within a team or group context.

These skills are trainable, and although research on skills training is still in its infancy, a variety of strategies and techniques exist to facilitate skill acquisition. Training of skills should proceed in a supervised environment and in a manner that is appropriate to an athlete's developmental level. Typically, developing skills requires investment and practice as well. Some techniques, particularly unfamiliar ones such as relaxation and meditation, may appear to be quite foreign and may frustrate beginners. But with repeated practice, they can be nurtured, with resulting improvement in performance.

Finally, although training of skills may be undertaken in a group or team format, the effective inculcation of psychological skills proceeds best at an individual level. Although some skill sets may be necessary for a class of athletes or team in general, individuals will often vary in the style or degree to which they implement a skill. In the end, efforts should be made to match skills training programs to individual athletes' strengths, weaknesses, desires, and capabilities. Attention to each athlete's uniqueness is ultimately the most important factor in enhancing his or her performance.

See Also the Following Articles

Anxiety and Optimal Athletic Performance ■ Arousal in Sport ■ Attention and Concentration Training in Sport ■ Goal Setting and Achievement Motivation in Sport ■ Group Dynamics in Sport ■ Intrinsic and Extrinsic Motivation in Sport

Further Reading

Hanin, Y. L. (1997). Emotions and athletic performance: Individualized zones of optimal functioning model. *European Yearbook of Sport Psychology*, *1*, 29–72.

Hanin, Y. L. (2000). *Emotions in sport*. Champaign, IL: Human Kinetics.

Mahoney, M. J. (1989). Sport psychology. In I. S. Cohen (Ed.), *The G. Stanley Hall lecture series* (Vol. 9, pp. 97–134). Washington, DC: American Psychological Association.

Mahoney, M. J. (2002). *Relaxing, reaching peaks, and learning process*. Plainfield, IL: Kinder Path Press.

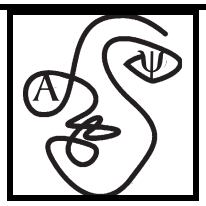
Mahoney, M. J. (2003). *Constructive psychotherapy: A practical guide*. New York: Guilford.

Moran, A. P. (1996). *The psychology of concentration in sport*. London: Taylor & Francis.

Orlick, T. (1990). *In pursuit of excellence: How to win in sport and life through mental training*. Champaign, IL: Leisure Press.

Singer, R. N., Hausenblas, H. A., & Janelle, C. M. (Eds.). (2001). *Handbook of sport psychology* (2nd ed.). New York: John Wiley.

Van Raalte, J. L., & Brewer, B. W. (Eds.). (2002). *Exploring sport and exercise psychology* (2nd ed.). Washington, DC: American Psychological Association.



Psychology and the Law, Overview

Roy S. Malpass
University of Texas at El Paso, El Paso, Texas, USA

Jane Goodman-Delahunty
University of New South Wales, Sydney, Australia

1. Introduction
2. Important Attributes of the Field
3. Articles in This Section
   Further Reading

GLOSSARY

adjudicative competence Functional legal abilities that a criminal defendant must have in sufficient degree in order for the adjudicatory process to go forward. Also called competence to stand trial, competence to proceed, and fitness to proceed.

best interests of the child doctrine The philosophy, now incorporated into law in all 50 states, that in making decisions concerning custodial placement and parenting plans, the interests of the children should supersede those of the parents and/or others involved in the dispute.

blank lineup (target absent lineup) A lineup in which all the members are known to be innocent.

cognitive interview Forensic tool that comprises a series of memory retrieval techniques designed to increase the amount of information that can be obtained from a witness.

direct questions Yes/no or short-answer questions that probe for particular details about an experienced or witnessed event; questions can range from nonleading to suggestive, depending on how the question is phrased and the type of information included in the question.

effective size The number of plausible members in a lineup.

encoding specificity principle States that if cues that were present at encoding are also present at retrieval, recall of stored information is more likely.

estimator variables Factors in eyewitness situations whose influence can only be estimated and are not under the control of the justice system.

ethical interviewing A police interview technique, based upon three principles, designed to encourage suspects to tell their side of the story. The use of deceit is forbidden, and psychological pressure is not used.

evaluation of comparative custodial suitability A forensic psychological evaluation in which the primary focus is the relative strengths and deficiencies in the ability of parents to parent the specific children whose custodial placement is at issue.

evidence-driven deliberation Deliberation in which jurors engage in general discussion of evidence in the case prior to voting, in contrast to verdict-driven deliberation, in which voting typically precedes evidence discussion.

forensic assessment instrument (FAI) An assessment test, tool, or device designed to measure or to inform a clinical judgment about specific functional legal abilities required in a given legal context.

forensic mental health assessment The evaluation of defendants or litigants, usually conducted by mental health professionals such as psychologists, psychiatrists, or social workers, that is performed to assist the court in making better-informed decisions or to help the attorney present evidence relevant to his or her case.

free-recall questions Vague, general prompts requesting for a narrative description of a prior events; prompts provide minimal cues or information about what should be recalled.

lineup bias A lineup constructed in a manner that leads the suspect to stand out from the other lineup members.

mental state at the time of an offense (MSO) A general term inclusive of various legally recognized mental states

©2004 Elsevier Inc.
All rights reserved.

(e.g., insanity, diminished capacity, automatism) that provide the bases for a defense to culpability for criminal behavior.

mock witness A person who is not a witness to a crime but who is asked to identify a perpetrator from a lineup based on another source of information, such as a verbal description.

post-event (mis)information Events that occur after a crime that influence eyewitness memories of the crime.

Reid technique A police interrogation technique, consisting of nine steps, designed to encourage reluctant suspects to talk during police interviews. The technique allows interrogators to use deceit and to build up psychological pressure.

restorative justice Legal proceedings that aim to make the victim and community whole following an offender's conduct, and to reintegrate the offender into the community, if appropriate.

scientific jury selection The use of scientific techniques by jury consultants to determine the individual demographic and attitudinal characteristics that are linked to views of a case. Techniques include community surveys, mock trials, focus groups, and in-court selection.

script A general knowledge structure that represents information from a class of similar events; a description of the typical, common features of repeated events.

showup The presentation of a single suspect to a witness in order to determine if that suspect is the perpetrator.

source amnesia The inability to recall the origin of the information that comes to mind when attempting to recall an event.

story model Theory of juror decision making in which jurors arrange evidence in the form of a narrative or story, and match the story to the closest verdict alternative.

structured interviews Scripted interview prompts and questions that are designed to obtain as much narrative information from children and be as minimally leading as possible.

suggestibility The act or process of impressing something (an idea, attitude, or desired action) on the mind of another.

system variables Factors that may influence eyewitness memory and that are under the control of the criminal justice system.

therapeutic jurisprudence Legal practices intended as therapeutic interventions for the benefit of the litigant.

therapeutic role Within the context of custody evaluations, perspective and actions that are consistent with one's obligation to assist those with whom one is professionally interacting.

verbal overshadowing effect When the act of verbally describing a perpetrator decreases the witness's ability to accurately identify a perpetrator from a lineup.

voir dire Questioning process during jury selection.

Psychology and law, or forensic psychology, is a branch of applied psychology, with three primary dimensions: research, practice, and policy implications. Major accomplishments in psychology and law are reviewed in this article, drawing on empirical research by experimental and clinical psychologists with several distinct forensic populations, e.g., juveniles, offenders, witnesses, and jurors. Experimental psychological contributions on detection of deception, eyewitness memory, and jury decision making are outlined. Clinical research and practices regarding forensic mental health assessments are described, e.g., of legal competency in criminal, civil, and family law cases. Roles of psychologists in therapeutic versus forensic settings are distinguished, and standards for providing relevant and reliable expert evidence are described.

1. INTRODUCTION

> If experimental psychology is to enter into its period of practical service, it cannot be a question of simply using the ready-made results for ends which were not in view during the experiments. What is needed is to adjust research to the practical problems themselves and thus, for instance, when education is in question, to start psychological experiments directly from educational problems. Applied Psychology will then become an independent experimental science which stands related to the ordinary experimental psychology as engineering to physics.
>
> The time for such Applied Psychology is surely near, and work has been started from most various sides. Those fields of practical life which come first in question may be said to be education, medicine, art, economics and law.

This bit of advocacy was written by Hugo Muensterberg in 1907, and makes clear that both the vision and the empirical research program on which the field is founded were present a century ago. Research at the time, and much of the work since, has taken place in the areas of perception and memory, as well as abilities, intellectual performance and social relations, areas in the list of problems related by Muensterberg. All of these areas are treated in the present section of this encyclopedia, and while they would all be recognized by turn-of-the-19th-century psychologists, their elaboration since that time has been considerable. Legal psychology has grown in many ways, and has multiple foci.

2. IMPORTANT ATTRIBUTES OF THE FIELD

There are certain attributes of psychology that set it apart from other fields with relationships with law

(e.g., social work, psychiatry). These attributes form common themes across the articles in this section. The field of psychology and law displays an interesting diversity of approaches to its application, which reflects the full diversity of the underlying field. Its approach is at base scientific, and in the analysis of policy, practice, and evaluation, the goals are theoretical and empirical clarity. The orientation is, for the most part, one of engaging problems at a practical level, using available theory and new research to develop pragmatic responses tailored to the demands of situation and context—just as Muensterberg foresaw.

The field of psychology and law exhibits a distinctive scientific character. The field is committed to testing standards that are based in measurement theory and developed according to scientific standards that distinguish this professional enterprise from that of others, e.g., social workers, psychiatrists. Validated measures and tests exist for some topics (e.g., competence to stand trial and competence to consent to treatment), but are not yet developed for others. The cutting-edge developments in the field are precisely the formation of new tests and measures, new empirically validated techniques and procedures, and policies founded in empirical research.

The potential contributions of psychology to the various demands of law are by no means fully realized. As psychologists embark on analysis of many domains in the psychology and law interface, alternative approaches and explanatory hypotheses are framed and tested using multiple data sources—a part of the fields of history of philosophy of science and methodology.

Psychology has a special role of assisting courts in making retrospective assessments of what occurred and in identifying lies and distortions (i.e., in serving justice). The contexts in which these needs appear are vast. Witnesses to events of all kinds, from bank robberies to traffic accidents, are relied upon to provide accurate information for legal processes. Those alleging the occurrence of a crime, either as a witness or as a victim, are evaluated for accuracy and for incentives to distort. Report of memory for events is a large part of many fact-finding activities in law. Psychology's special role emerges from the fact that records of events and the accounts of them come from human memory and the perceptual and cognitive processes that both record and report them—both selectively. The report of a police officer, for example, comes from memory, and even when it is written contemporarily with the event, the selectivity typical of human attention requires understanding and interpretation, especially if the officer's report conflicts with the reports of others. Psychology is uniquely positioned to address memory for events and other forms of experience, as well as intentional distortion and deception. No other field of study is so focused on these processes.

Memory for narrative integration of information, making judgments, and studying the effects of personal history, attitudes, and belief systems are all aspects of legal processes. Some take a central role and assume great importance. Jury decision making is a strong traditional area of psychology and law. The integration of new information into existing attitude and belief systems in the context of whatever constraints are placed on judgment by the court is an important area of understanding for scholars to pursue. Research in this area has the potential for important contributions to our understanding of why long-term trends of injustice (e.g., racial discrimination in all areas of opportunity and legal protection, gender inequality in ownership and employment) were able to continue and how and why change occurs.

Criminal suspects and litigants have incentives to fabricate, mislead, exaggerate, and confabulate. These incentives and the difficulties they present to fact finders extend through many areas, including forensic mental health assessment. As noted by Heilbrun and Landers in their article on assessment in this section,

> Some individuals evaluated through FMHA respond to questions openly and honestly. Others exaggerate or fabricate symptoms and problems. Yet others minimize or deny symptoms that are genuinely experienced. Response style is a term used to describe how an individual reports his or her own symptoms, behaviors, and problems, and is used to classify an individual's responding as (1) reliable, (2) exaggerated/malingering, (3) defensive, or (4) uncooperative. Self-reported information from an individual who is not reliable will have significant inaccuracies. This principle emphasizes the value of third party information (records and interviews of collateral observers) in both detecting and correcting these inaccuracies.

The role of psychologists in a legal context is often structured very differently from that of psychologists in other contexts. For example, there is an important distinction between therapeutic and forensic roles. Ten differences elaborated by Heilbrun include (1) the purpose for which the evaluation is conducted, (2) the nature of the relationship between the evaluator and the individual being evaluated, (3) the need for a

formal notification of purpose, (4) who is being served, (5) the nature of the standard(s) applied in the evaluation, (6) the sources of data considered, (7) the response style of the individual being evaluated, (8) the need to clarify the reasoning and limits of knowledge, (9) the nature of the written report, and (10) the expectation of expert testimony.

Heilbrun and Lander remind us that it is only in the past 50 years that courts have been using psychologists as opposed to other mental health professionals to provide information on assessments. Recognition of the role of psychology is still lagging in some communities, and many courts still require a medical degree from the professionals whose opinions they seek. It is only in the last 25 years that expert testimony by experimental psychologists has been widely used in court in matters involving memory and eyewitness identification.

The role of expert witness is quite different from the role of researcher and teacher. An important example is ultimate opinion testimony. A researcher specializing in eyewitness identification will rarely speak with the witness(s) who made the identification, and so is sometimes asked how he or she can talk about the accuracy of the eyewitness identification without examining the witness. The answer is straightforward and contains an important constraint on the role of an expert in contrast to the fact finders in the legal system (e.g., juries, judges). The important factors influencing eyewitness identification are either facts elicited in depositions, hearings, or in trial or facts about the construction and administration of the eyewitness identification materials and process. These are all external to the witness him- or herself. It is a misconception (and maybe a stereotype) that experimental psychologists testifying as expert witnesses on eyewitness testimony make psychological assessments of the witness. The question also contains a trap and another misconception. Eyewitness experts do not offer conclusions about the accuracy of the eyewitness identification. On the contrary, they discuss the events and conditions that are known to influence the accuracy of identification. It remains the province of the jury (or other fact finder) to consider these factors in their own consideration and judgment about the probable accuracy of the identification. Offering an opinion about the accuracy of the identification is known as ultimate opinion testimony, in that it offers an ultimate opinion about an issue that is to be decided in the trial process. The role of the expert in this instance is to inform the jury so that in their evaluation of the evidence they can consider the expert's information—a kind of teaching role for the expert. It is important to keep the decision making in the hands of those placed in that role by the legal process; this clearly structures the role of experts in court and related proceedings.

3. ARTICLES IN THIS SECTION

Nowadays, forensic psychologists are frequently consulted by courts or lawyers to provide information about a party's mental state or a party's decision-making ability or capacity, although some judges and other court personnel remain confused about the differences in training and expertise offered by a psychologist, a psychiatrist, a social worker, or other mental health professionals. The articles by Poythress, Heilbrun and Lander, and Martindale and Gould serve to clarify the nature and scope of empirically based contributions that psychologists can make by providing forensic mental health evaluations in civil and criminal cases.

The extent to which an individual is personally responsible for his or her conduct is critical in many situations involving both criminal and civil legal issues. For example, a defendant who engages in behavior that is illegal, but who is unaware of what he or she is doing, or who does not appreciate the significance or consequences of these actions may be entitled to a defense against criminal charges on this basis. Similarly, someone who is easily influenced because of a learning disability or some other psychological impairment can not be expected to take charge of his or her own affairs to the same degree as someone without any mental impairment, and once his or her disability is established, may require assistance in the form of a legal representative or guardian. Expert testimony or a report from a mental health professional such as a psychologist may assist the courts in understanding the cognitive limitations of any impairment, temporary or permanent, as it bears on a legal issue. Psychologists who are trained in forensic assessment and evaluation can help courts to resolve such legal issues.

Defenses to criminal charges premised on a mental disease or disorder are often unpopular in the community because of common erroneous assumptions about the frequency and consequences of such a plea. Research by Silver *et al.* on the public perception of the number of defendants who plead not guilty by reason of insanity shows the number is often exaggerated, and moreover, that the public wrongfully assumes that defendants who enter such a plea are released into the community. In fact, the insanity defense, which applies

only to the defendant's state of mind at the time the alleged offense was committed, is used in fewer than 1% of all criminal cases, and when proposed it is unsuccessful 74% of the time. As Perlin points out, the consequence of entering a plea of insanity is often a period of incarceration in a psychiatric hospital longer than the offender would spend in prison if convicted of the crime charged.

Other defenses proffered on grounds of diminished capacity or limited mental responsibility are similarly unpopular and controversial, as they, too, are often thought to permit culpable offenders to circumvent punishment for their transgressions. Expert assessment and expert testimony is common in these cases to explain how varied mental or neurological conditions may produce situations in which the conduct of the accused is unintended, e.g., because it was involuntary, perhaps a form of automatism, caused by events such as a head injury or a condition such as sleepwalking, hypoglycemia, epilepsy, consumption of alcohol or drugs, or dissociation. Certain mental conditions or disabilities may exonerate a defendant from criminal responsibility if the offender lacks the capacity to form the necessary criminal intent, e.g., defenses that the offender was suffering from battered woman syndrome, post-partum depression, sexual addiction, urban survival syndrome, false memory syndrome, post-traumatic stress syndrome, multiple personality disorder, or schizophrenia.

3.1. Legal Competency

The article on legal competency by Poythress delineates the types of decision-making abilities or competencies at discrete moments in the civil and criminal legal process about which courts most frequently seek advice from psychologists trained in forensic assessment and evaluation. The accompanying tables and charts succinctly illustrate the legal standards or issues in the cases and the related psychological competencies that bear on them. Three elements that must be addressed by a forensic psychologist in each case involving an assessment of fitness or competency are (1) the presence of a mental condition that causes an impairment, (2) specifications of the functional impairment, and (3) the distinct legal abilities key to the case that are affected.

The article begins by orienting practitioners to the context of the task by outlining a framework within the law that psychologists are well advised to understand in order to offer more useful information to the courts. A helpful conceptual tool for practitioners and psycholegal researchers is a table listing the five aspects of legal competencies identified by Grisso, as it sets forth those aspects of the task most clearly within the province of the forensic psychologist, and those aspects most clearly within the province of the fact finder (judge or jury). In particular, it is important to note that the legal standards or tests for fitness vary from one circumstance to the next. Accordingly, the psychologist must appropriately tailor the information critical to the fact finders in applying the legal standard appropriate to a given case. Many of the competence tests and standards pertinent to civil cases have received less research attention than those pertinent in criminal cases, i.e., criteria relevant to determinations of fitness to make decisions about guardianship, treatment, research participation, and testamentary capacity.

A point of significance is the limited utility of a diagnosis in many legal proceedings. Although the goal or end point of many clinical mental health inquiries is a diagnosis, legally, the diagnosis itself is of little value to a court charged with determining whether someone meets the criteria for a defense of insanity or otherwise lacks fitness to proceed to trial. The fact finder seeks a functional description or understanding of the nature and scope of an individual's impairment from the consulting expert. A diagnosis by a mental health professional of a mentally disabling condition or disorder does not guarantee that the applicable legal test will be met, but the diagnosis may be of some relevance. Severe mental disturbance or lack of a diagnosis neither guarantees nor is fatal to a claim of incompetence or insanity.

The terms insanity and incompetence are legal, not psychological, concepts. The legal test for insanity varies from one jurisdiction or community to another, although most tests incorporate the offender's cognitive awareness of the consequences of the conduct at issue. Tests that address some of the "irresistible impulse" standards have in some cases devolved into a question as to whether the accused would engage in the same conduct were there a policeman present. In some states or countries, the standard of "guilty but mentally ill" has been adopted as a variant of "not guilty by reason of insanity." In other communities, legal policy reforms have established that voluntary conduct that produces diminished capacities, such as drinking alcohol or taking drugs, obviates resort to this defense. Finkel points out that policies along these lines conform with the results of studies of juror decision making in insanity cases, which revealed that jurors often take into account the offender's capacity

to make responsible choices and whether the offender was negligent or reckless in bringing about his or her mental disability.

In his article, Poythress adroitly sets controversial topics such as the insanity defense in context by pointing out the infrequency of its use and its low success rate. He also emphasizes the importance of and demand for many other fitness determinations by psychologists, such as fitness to stand trial, or when the death penalty applies, fitness to be executed. Although it is difficult to obtain precise figures of the number of evaluations prepared annually on issues of fitness or competence, Costanzo shows that approximately 5% of all criminal defendants are assessed for competence to stand trial. Thus, Zapf and Roesch provide estimates that between 25,000 and 39,000 psychological evaluations for competence to stand trial are performed annually in the United States, making this one of the most frequently requested psychological services. Of the group of offenders evaluated for this purpose, Melton *et al*. indicate that a relatively small proportion, 12%, is found incompetent.

This article provides an excellent overview of methodological issues facing professionals who are bound to offer evidence-based conclusions to the court. The article includes practical advice to practitioners, such as avoiding controversy as an expert by refusing to prepare a report that addresses more than one type of fitness. Similarly, psychologists are advised to avoid making any statement about the "ultimate legal issue." As Poythress points out, to do so mixes moral and professional standards.

Poythress also illustrates the significance of a clear understanding of the relevant comparative group in interpreting the test outcomes of standardized test instruments. Over the past decade, more test instruments designed to address legal competence and fitness have been developed and studied, thus the dearth of available forensic tests is decreasing. One of the drawbacks in developing evidence-based instruments is finding appropriate control groups and obtaining permission to conduct randomized trials on relevant forensic populations to produce comparative data and appropriate test norms.

3.2. Forensic Mental Health Assessment in Legal Contexts

The article by Heilbrun and Lander on forensic mental health in legal contexts complements the work of Poythress by providing a more hands-on, practically oriented discussion. Heilbrun and Lander include step-by-step guidance on how to conduct a forensic mental health assessment, including advice on how to structure a report intended for use by the court. Heilbrun and Lander guide the forensic practitioner through 29 steps for conducting a forensic mental health assessment, starting with the initial consultation and ending with expert testimony. This article leads practitioners through substantive considerations at each of three major steps in conducting forensic assessment, namely, preparation, data collection, and data interpretation. For each step, the authors itemize the standards or guidelines that provide parameters for sound practice. Thus, this article is of value in outlining the various pertinent authorities: legal, ethical, scientific, and practical or clinical.

One of the vital themes developed for sound forensic practice is the importance of gathering information from multiple sources to permit more extensive scrutiny of the reliability of the information gathered in the context of litigation. Heilbrun and Lander state the following:

> A single source of information is usually less accurate than multiple sources considered together. This is particularly applicable in forensic mental health assessment, as some of the information sources can be particularly distorted. A defendant might be inclined to exaggerate or minimize symptoms or capacities, depending on the possible implications of each in litigation.
>
> Others exaggerate or fabricate symptoms and problems. Yet others minimize or deny symptoms that are genuinely experienced. Response style is a term used to describe how an individual reports his or her own symptoms, behaviors, and problems, and is used to classify an individual's responding as (1) reliable, (2) exaggerated/malingering, (3) defensive, or (4) uncooperative. Self-reported information from an individual who is not reliable will have significant inaccuracies. This principle emphasizes the value of third party information (records and interviews of collateral observers) in both detecting and correcting these inaccuracies.

Other researchers, such as Rogers, have developed forensic instruments to assist practitioners in discerning distorted responses from litigants who seek to appear better or worse than they are in reality.

A second major theme of forensic mental health assessment in legal contexts, developed by Poythress, Heilbrun, and Lander, is acknowledgement of the

distinction between psychologists who are retained to perform a therapeutic role, e.g., by treating offenders, and those who are retained to perform an evaluative forensic role. The full nature of the distinction is appreciated when one considers the ten differences between these roles elaborated by Heilbrun, outlined previously in this article.

Psychologists are often consulted to assess parental competence in the context of a marital dissolution, where decisions about the placement of children in the custody of one, both, or neither of the parents must be made by a court. One parent may be mentally ill, or there may be allegations of physical or emotional abuse or neglect by a parent; in such cases, the courts must determine whether that parent is unfit to raise the children. More frequently, courts must grapple with the question of which type of custodial arrangement will promote the child's best interests.

Psychologists who are well versed in the research on psychological effects on children of divorce and custodial arrangements can play a significant role in assisting courts with determinations about parental fitness. Family life has traditionally held a sacrosanct position within most legal systems. Over time, social and cultural factors regarding who has worked outside of the home and who has been a primary caregiver for a child have shifted the status and rights of children, women, mothers, and fathers. As Krauss and Sales point out, definitions of the family have undergone a dramatic change, particularly regarding the nature of nuclear and extended families. Hetherington *et al.* show that about one-half of all Australian and North American children born in the 1990s will be raised by a solo parent, and many children have no siblings. The rise of "no fault" divorce laws has contributed to these changes and has removed much of the stigma previously associated with marital dissolution. However, in certain cultures, remaining married to a spouse who takes on a new partner provides more benefits and social advantages to the ousted wife and children than does the option of divorce.

Along with changes to the definition of family, the concept of the psychological parent has developed. Accordingly, the rights of grandparents, step-parents, gay parents, and other domestic partnerships have influenced trends in custodial determinations. The extent to which children are consulted regarding their preferences, and the age at which this occurs, varies from one community to another.

The contributions of and participation of psychologists as researchers and practitioners in issues relating to marriage, divorce, and child custody have been welcomed by legal practitioners. Bausermann demonstrates how research has contributed to our understanding of circumstances that promote healthy adjustment in children, and can assist in assessing custodial options such as sole versus joint custody. Eight studies that compared children's adjustment in intact families versus joint custodial arrangements showed no adjustment differences in the children. Data from 33 studies indicated that (1) no causal relationship has emerged between the type of custodial arrangements and children's adjustment; (2) when the parents are in extreme conflict, children experience adjustment difficulties no matter what the custodial arrangement; (3) more parental conflict ensues between ex-partners when one parent retains sole custody than when custody is joint legal and physical; and (4) the ongoing involvement of the father with the child is predicted by the relationship between the parents, not by the relationship between the father and the child.

3.3. Child Custody

The article on child custody by Martinson and Gould traces the four major doctrines relied upon by courts in making custody decisions: (1) the "tender years" doctrine for children under the age of 7, favoring maternal custody; (2) the "psychological parent" rule, favoring the primary caregiver; (3) the "least detrimental alternative," favoring the stability and ongoing relationships in the child's life; and (4) the "best interests of the child." Although the latter is the prevailing standard, it has repeatedly been criticized as too vague, e.g., regarding the time frame to take into consideration when it is applied.

This article also outlines the key features of ethical practice standards, such as the American Psychological Association Guidelines on Custody Evaluations. These guidelines assume the legal criterion is the best interests of the child. In line with the distinction between functional and diagnostic criteria noted earlier, the guidelines emphasize that evaluators must avoid focusing on the psychopathology of parents except to the extent that it affects parenting skills and fitness. The guidelines echo the need to use multiple methods of data gathering, such as observations, interviews, contacting third parties or collaterals courses, obtaining school reports and records, as well as formal psychological assessments, in order to offer a more reliable recommendation to the courts.

Because the stakes are high in custody disputes, practitioners are advised to take steps to uncover efforts by the parties undergoing evaluation to exaggerate positive information or minimize negative information. The principle of substantiating one's opinion with evidence to back it up is critical to avoid recourse to clinical opinions that are more susceptible to attack. Many of the recommendations for successful forensic consultation in child custody cases can be applied to other areas of forensic practice.

Commentators have pointed out that one consequence of the implementation of no fault divorce laws is the greater likelihood of resorting to mediation to resolve issues of custody. In many jurisdictions, parties litigating in family court must participate in mandatory mediation or family conferences prior to attending any court hearing, and in other jurisdictions, referral to mediation is routine. More custody cases are resolved via mediation than by trial. Research on uses of mediation in family cases is more extensive than in other areas of the law. Sales and Beck argue that mediation may produce more losses for mothers and gains for fathers and can have both positive and negative effects. Success of mediation has more to do with the readiness of the parties for the procedure. The best candidates for mediation are those who are cooperative in resolving conflicts. There is no evidence that the psychological functioning of parents enmeshed in conflict can be enhanced by mediation, or that mediation can change the relationship patterns of parents who are angry, volatile, disengaged, or withdrawn. Researchers also caution against generalizing from measures of satisfaction with mediation.

The approaches to evaluation and assessment outlined by Poythress, Heilbrun, Lander and Martindale and Gould establish core principles than can be applied to all forms of forensic assessment. Greenberg and Brodsky discuss their application to civil damages and assessment of the nature and scope of injuries following an accident, and Foote and Goodman-Delahunty discuss their application following a statutory violation, such as employment discrimination.

Just as children have been given greater and more careful attention in many legal processes than in the past, more recently their testimony has achieved greater acceptance. They have been seen more frequently as victims of crimes and therefore as witnesses. Many nations have experienced an increased focus on children and their welfare in a wide range of legal contexts. As divorce rates have increased, the question of how to provide for the children has come more into focus. In that context alone, there has been an increased concern for the abilities of children to know their own minds, to give accurate testimony, to withstand the pressures of family members, and to participate in the legal process. Although many of these concerns focused on protecting children from becoming victims of their situation, many children do indeed become victims, often within the family. Various kinds of abuse are observed, from neglect, to beating, to sexual abuse. Abuse that in the past would not have come to light in public or in legal proceedings came to the attention of professionals working for public agencies charged with overseeing (among other things) child welfare. As these events are reported, questions inevitably arise regarding the abilities of children to remember the important events, report their memory, differentiate between what they experience directly from what they have been told, and withstand the pressures of testimony and cross-examination in the presence of strangers and the imposing atmosphere of the courtroom. There are also concerns about the appropriate means of interviewing children to get the most reliable information with the least cost to the child.

3.4. Children's Testimony

Quas and Beck divide the research questions that have been asked in this area into three major categories: characteristics of the child, characteristics of the to-be-remembered event, and characteristics of the interview. Age is a rough way of indexing the cognitive and social changes that occur with psychological development. Many changes occur as children grow and age: their general knowledge, sense of what is important, and ability to interpret and remember events all change in important ways. Other factors (e.g., intelligence, temperament, emotional reactivity) are also important influences on the child's memory and ability to provide reliable testimony. The authors point out that it is difficult to apply knowledge of these factors to make judgments about the validity of a given child's report. However, such information can assist in deciding how best to interview the child, and how to protect the child during the process. The recent move by the Home Office in the United Kingdom to identify and moderate treatment of vulnerable witnesses is of particular interest in this connection.

Event characteristics are of interest here just as they are in other areas of eyewitness memory. Quas and Beck review the major findings from research with children. Emotional events, for example, are remembered better than neutral events, although peripheral

aspects of an emotional event may not be well remembered. But again, applying this knowledge is not always straightforward. What is central to a child may not be what is central to an adult, and as with adults, children vary with regard to what aspects of an event they consider central or peripheral. Like adults, children learn generalized accounts of common, repeated events. But more so than adults, they fail to differentiate between the common characteristics of an event and the unique ones, often failing to report unique events and common, expected events that did not occur in the instance of interest.

The characteristics of both interviews that lead to erroneous reports and those that lead to more accurate and complete reports have been the subject of a very large portion of research on children's testimony. To some degree this was driven by the nature of the court cases that were very much in the public eye in both North America and Europe. The conclusions reached for children are very similar to those developed by Vrij for interviewing adults. Free recall questions are fairly accurate (although they may be incomplete), and suggestive questions lead to errors, in the direction of adopting the suggested information.

3.5. Interviewing and Interrogation

Obtaining information from persons associated with a criminal event (and certain other events such as food poisoning, accidents, and workplace events) is a process that is both important and fraught with possibilities for error. Some errors are intentional—e.g., withholding self-incriminating information—while others are mired in the complexities of retrieving precisely accurate memories. The difference between these cases is the difference between interrogations and interviews. This and the related topic of detecting deceitful testimony are treated in the article by Vrij on interviewing and interrogation.

Posing suggestive questions to victims and other witnesses risks the integrity of the resulting information. When witnesses are asked leading questions containing one possible answer to the question, e.g., "Was the car red?", when the answer is "Yes" it is not clear whether it is because the car was in fact red or because the witness has no memory of this fact but answers "Yes" because of a judgment that the questioner probably knows the right answer and is looking for confirmation. Alternatively, the suggestion may actually modify the witness's memory. On subsequent occasions, the witness may confidently and spontaneously answer "Red" and experience the attribute of "red" associated with that object and event, firmly believing it was red.

One alternative to asking leading questions is to ask questions containing multiple choices, but these are still limiting and leading, albeit in a broader sense. The best practice is the use of open questions, often a number of open questions, sequentially. For example, "Tell me everything you can remember about the car," followed by "What was the color of the car?" if color is not mentioned. In general, it is better to begin with the more open questions, probing with more open questions, and ending with specific questions (leading, in the end) to elicit specific responses. However, if the witness cannot come up with "Red" in response to "What color was the car?", asking "Was it red?" yields interpretable information only if the answer is "No." These topics are discussed in more detail by Vrij.

For more than 20 years, techniques based on context reinstatement have been known to enhance the recall of information, even after significant time delays. The most commonly used of these techniques is the cognitive interview (CI), as developed by Fisher. CI techniques are based on the premise that the witness has the information wanted by the interviewer, that the information is organized by the witness according to meaningful criteria applied by the witness at the time of encoding or organizing the information for retention, and that the retrieval of this information must be organized by the witness (not the interviewer) in ways that fit the organizational criteria imposed on it by the witness in the first place. This set of assumptions requires that much of the social interaction—as well as the cognitive interaction—about the recall process be in the hands of the witness. Witnesses are therefore encouraged to take an active part in the interview, to recall and report everything, and to return to previously discussed topics if additional details have been recalled. Witnesses are also encouraged to report the events under investigation from the perspective of other actors present during the events. Research indicates that context reinstatement techniques greatly increase the gross amount of memory reports. The ratio of correct to incorrect items recalled does not change, however, so while the amount of material produced is greatly increased, its efficiency at differentiating correct from false reports is unchanged. Nevertheless, having much more to work with is a great advantage to investigators.

The Reid technique is the single most popular approach to the interrogation of uncooperative suspects. It is

used to obtain confessions from people whom investigators believe are guilty. Of course, the detection of guilt or innocence is well known to be a judgment prone to error, and the confession is meant to establish guilt, so the technique is to some degree circular. Its adherents defend it as being nearly perfect, producing confessions from the guilty but not from the innocent. However, there is an increasing catalog of cases in which Reid-like techniques are associated with false confessions and wrongful conviction. The technique has achieved a negative reputation in psychological research circles because of certain components of the technique, such as altering the suspect's decision-making environment by lying about the nature and strength of the evidence against him, by painting the picture of his situation as hopeless, and creating a (false) understanding that confession is the only way he can obtain any leniency (something usually outside the investigators control). Vrij's article reviews the Reid technique and provides an extended analysis of its difficulties.

Vrij offers as an alternative a technique used increasingly in the United Kingdom as a response to a series of notorious British miscarriages of justice involving false confessions. The technique involves the extensive use of detailed knowledge of the crime by the investigating officers and the skillful use of questioning techniques.

False confessions have received a great deal of attention in the popular press in the last few years, and many nations have their own catalog of prominent cases. Vrij distinguishes three types of false confessions: voluntary, coerced-compliant, and coerced-internalized. Voluntary false confessions are given with no special effort from investigators. The famous case of Henry Lee Lucas, who is estimated to have confessed to more than 600 murders is offered as an example by Gudjonson. Coerced-compliant confessions are those that result from the pressures and incentives of an interrogation. Coerced-internalized false confessions occur when people become convinced that they actually did commit the crime even though they have no memory of the event. Analysis of these forms of confession is provided, and means of detecting false confessions are discussed.

Law enforcement personnel often display great confidence in their ability to detect deceit in suspects they are interviewing. The bad news is that in controlled studies, people are able to detect truthfulness a bit above the level of chance expectation but detect lies somewhat below the level of chance expectation. Studies using law enforcement personnel, as compared with college students, find similar accuracy rates. Some professionals (e.g., U.S. Secret Service officers and others with a special interest in deception) perform better than chance levels. Vrij reviews a number of common pitfalls in the deceit detection process, discusses cultural differences in social interaction norms that interfere in the process, and debunks some popular beliefs about the behavior of liars. Techniques for improving deceit detection are reviewed, e.g., attention to the content of a witness's speech leads to more effective assessments of deceit than does attention to non-verbal behavior.

Another form of error, although rarely intentional, is that made by eyewitnesses when they are asked if the perpetrator is among the persons shown to them in a "lineup." The prevalence rate of such errors is not altogether clear. However, the problem has persistently attracted attention for more than 100 years and is not easy to dismiss. The most recent incarnation of attention to this problem is found in the work of the Innocence Project, which as of April 2004 has documented 143 exonerations of wrongfully convicted persons. Eyewitness identification errors are the single most important contributor to these wrongful convictions. The malleability of memory is an important contributor. New information can be incorporated into old memories, and the source of memories is often forgotten. It is often not possible to tease out the origins of what appear to be intact and accurate memories. In addition, it is not uncommon for aspects of a typical event to be incorporated into memory for a specific event in which certain of the typical elements were not present.

3.6. Eyewitness Testimony

The misinformation effect, discussed by Levine and Loftus in their article on eyewitness testimony, is an important example. They show how the wording of a question can lead to the intrusion of non-existent elements into reports of memory. Similarly, the strength of the verb used to characterize and event (e.g., "How fast were the cars going when they smashed/collided/bumped/contacted ... ?") affects the answer. Not surprisingly, the effects are stronger when the actual memory is weaker. And the passage of time only fosters integration of the new or modified elements into an altered but apparently unitary memory.

The authors review the relatively new and somewhat surprising findings on the subject of creation of new memories of personal events that in fact never

happened. Suggestions made by trusted others are important ways in which non-existent events can be created in memory. But imagining an event oneself can also result in the creation of memories, as the fact that the event was only imagined is itself a memory that may be forgotten.

The spurious relationship between confidence in the accuracy of a memory and its actual accuracy is briefly discussed. Belief in the validity of a confidence accuracy relationship is strong on the part of many people, some of whom will serve as jurors or fact finders in legal contexts. Even though confidently given accounts of events may be convincing, the confident manner of reporting is in itself no indication of the validity of the recollection. Only a very small relationship is found between confidence and accuracy. And to make matters worse, confidence is easily manipulated independently of accuracy, making it an even less useful factor.

The role of emotion in memory is an interesting area of study. Emotional events appear to be well remembered, but it appears that this is primarily for central events, or those that have important meaning for the observer. Peripheral events appear to be more poorly remembered. Studies of activation of brain areas and their relation to memories of emotional events can help to clarify our interpretation of the memory–emotion relationship.

The confrontation of a criminal by victims and other witnesses captures interest and imagination; when the accusing finger points at the perpetrator of a crime, it is the moment of justice. However, this is no simple imposition of justice at a personal level. Everything about the identification process has meaning for the outcome, and eyewitness identification is notoriously fallible. Because the identification and the preparations for it have a central role in a significant proportion of criminal cases, concern for identification procedures has attracted the attention of law enforcement agencies, non-governmental organizations concerned with justice, and the research community.

3.7. Eyewitness Identification

Tredoux, Meissner, Malpass, and Zimmerman introduce the topic of eyewitness identification and set forth some of the complexities of this area of applied research. The authors distinguish three stages in the eyewitness process: perceiving an event, storing and assimilating the information obtained, and remembering and acting upon the event. They review and discuss the ways in which factors related to these stages affect the result of the identification process.

Any identification is dependent on the amount and quality of the information obtained from the original event. This is affected by the amount of time the witness had to view the event and the perpetrator, distance from the event, and its visibility, including lighting and conditions that obscure vision. Apart from external conditions, aspects of the witness are also important at the initial stage: stress, fear, and the presence of a weapon all influence the information-gathering process.

Certain enduring characteristics of witnesses are related to their performance. Their age (at least at the extremes) has some influence. Alcohol and other drugs also have effects. The "race," sex, and occupation of witnesses have also been studied. While there is no general advantage of males or females in identification performance, there are some differences in memory for events that appear to be matters of attention rather then memory per se. While it is a popular belief that law enforcement officers will be better in facial identification, the small research literature on this topic does not support this idea. However, experience as a officer does relate to the quality of the descriptions they provide, both of persons and events. However, they also tend to misperceive innocent events as being criminal in nature. The race of the witness is unrelated to facial or event memory of reporting performance.

Attributes of the perpetrator also contribute to accuracy of identification, in interesting ways. First, disguises are effective, especially those that cover the most informative areas of the face. There are also distinctive patterns of facial appearance. Typical faces are more difficult to distinguish from each other, while nontypical faces are more distinctive and easier to remember. Some faces can be characterized as having specific characteristics (such as "criminality"); observers agree on this, even though it is unrelated to actual criminality. The race of the perpetrator alone does not contribute to accuracy of identification, but the combination of the race of the perpetrator and the witness does. "Cross-race" identifications are more frequently in error than "own-race" identifications.

An important part of the investigative process is getting information from witnesses, about both the event and the appearance of persons taking part in the event. Verbal descriptions of perpetrators are notoriously sparse, but there are techniques that assist first responders in obtaining more useful descriptions and information; some of these are described. Consequences

of asking witnesses to describe the faces of perpetrators include short-term reduction in the ability to recognize the face described.

The technology of pursuing the identification of suspects by eyewitnesses includes such procedures as searching mug shot collections, constructing composite portraits, and constructing and administering lineup and other identification procedures. These topics are all discussed in this article. More detailed attention is given to the construction, evaluation, and administration of lineups.

It is important for law enforcement organizations to keep collections of facial images so that fair lineups can be constructed for criminal investigation. The authors point out that these collections and the means of perusing them are behind the available technology. Examining mug shots is a facial recognition process. Facial recall, however, externalizes a facial image in memory to a representation that can be shared. Verbal descriptions are a form of face recall; another is the construction of a composite portrait. Of course, if the portrait is constructed by another person through the witness's verbal descriptions and then approval by recognition, it is not exactly facial recall. But when the witness uses composite construction computer programs, the work of the eyewitness is more constructive. These are moot points, however. There is no evidence that composite construction is a reliable means of producing high-quality likenesses from memory. There are other dangers in their use: first, many investigators believe in their usefulness, and second, many times a witness is given a copy of their composite to keep with them. Evidence suggests that repeated viewing of a composite changes memory for the original face.

Lineups should be constructed to be fair in two respects: the suspect should not stand out from the other members of the lineup, and the fillers—the five other members of a six-person lineup—should be meaningful alternatives to the suspect. Quantitative methods for measuring the fairness of a lineup have been developed and can readily be applied by law enforcement personnel who create them. The authors describe these techniques and show their importance for effective use of identification evidence in an investigation and prosecution.

Many law enforcement agencies have policies governing the admonitions that are read to eyewitnesses during the identification process. The authors argue that one of the more important problems in eyewitness identification is the tendency for eyewitnesses to believe that they must choose someone from the lineup they are shown. Admonitions of various kinds have been developed in an attempt to correct this tendency, and to orient the eyewitness to the important nature of the identification decision.

Alternative forms of lineup procedures have been proposed and researched in the last decade. The authors describe and evaluate a series of these, and describe what is known about the use of lineups involving other modes than visual sensory ones. Finally, the authors consider the beliefs and knowledge held by prospective jurors about eyewitness identification. This is information that feeds into the evaluation of the use of expert testimony in the evaluation of eyewitness identification. The balance between informing the jury and influencing the jury is an important one, and so it is important to know whether juries can benefit from expert assistance. Research indicates that lay jurors lack the requisite knowledge to appropriately evaluate eyewitness identification evidence.

3.8. Jury Decision Making

Opportunities for laypersons to participate in legal decision making exist in many jurisdictions, most commonly provided through some form of jury service, although there is considerable variability in jury systems and the scope of the jury role and the types of cases in which the public participate around the world. Albertson, Farley, and Hans report on developments in the jury decision making area. Jury behavior is one feature of the legal system that has been the subject of extensive psychological research for several decades, starting with the University of Chicago Jury Project. The research database on jury studies is so vast that gaining an overview is challenging. A recent meta-analysis of jury research published in an 18-year period (1977–1994) revealed that studies of criminal cases far outstripped those about civil jury decision making (89 vs 11%). Thus study also provided some insight into the most frequently studied topics of jury behavior: the impact of litigation strategies, the influence of different types of witnesses, and the effects of juror or litigant characteristics. Distinctive subtopics addressed most often by researchers emerged: criminal defense strategies, capital punishment and sentencing, characteristics of individual jurors, judicial instructions to jurors, general litigation strategies, witness performance, evaluation of evidence by jurors, and characteristics of offenders or parties.

Jury researchers have employed a diverse array of methods and approaches to gain insight into factors

that influence jury performance. These include archival studies of jury verdicts in actual cases, case analyses, community attitude surveys (e.g., via telephone), observations of trials and interviews of actual jurors, experimental simulations, and experimental or quasi-experimental designs using real juries.

One of the limiting factors in studying jury decision making is that the deliberations that take place among jurors at the conclusion of the trial are typically confidential, and the observation or recording of jury deliberations in many jurisdictions is prohibited. This policy has impeded researchers from gathering much information about the scope of jury-initiated misconduct in actual trials. Researchers who rely on post-trial interviews can not accurately gauge the extent to which juror recall is accurate and reliable. At times, researchers have employed "shadow jurors" who attend courtroom trials and then are interviewed about their observations to infer what the actual jurors might think of the evidence. Occasionally, extra alternate jurors have been empanelled and their deliberations have been videotaped and analyzed to gain insight into how the information they heard and witnessed is used in reaching a verdict.

Research involving real jurors in actual cases, including archival studies of jury verdicts, is relatively rare, accounting for approximately 10% of the jury studies conducted over an 18-year period. Because this method is not very informative about the decision-making process of the jury, many researchers have adopted a mock trial paradigm in which simulated juries receive information akin to that presented in an actual case. The materials may range in realism from brief written descriptions or case summaries to extensive videotaped trials in which attorneys present opening and closing arguments and question witnesses via direct and cross-examination, and jury instructions on the law are provided to the mock jurors. A benefit of this approach is the more extensive empirical control gained by the researcher, permitting causal inferences to be drawn. When mock jurors are used, the extent to which they are similar to actual or potential jurors is an issue. Although much of this research has been conducted with psychology undergraduates, the use of jury-eligible participants has steadily increased, enhancing the generalizability of these results.

The unit of analysis studies by jury researchers has varied. In some cases, individual jurors are the target of the studies, while in others, group deliberations are included, and the unit of analysis is the jury. The article by Albertson, Farley, and Hans carefully outlines the major findings to emerge from both types of studies, i.e., those that have focused on individual jurors and those that have addressed the jury as a group. Studies of the deliberation process have increased in the past 30 years, and they have examined the influence of features such as the size of the jury, the way in which differences of opinion between majority and minority groups are resolved, the influence of majority versus unanimity rules, and so forth.

The significance of distinctions between external validity and ecological validity have been highlighted by Bornstein and Vidmar, specifically in discussing whether research conditions in simulation studies include sufficient features of real trials to render them adequately representative of actual trials, permitting researchers to draw on research findings to make policy recommendations regarding reforms to enhance jury performance. Questions of normative criteria against which jury performance should be assessed have also been raised, particularly since research has shown that many judges are susceptible to the same biases or weaknesses as are many jurors. The authors suggested that profitable future avenues of research might include the influence of juror and jury anonymity, options for jurors to become more active during the trial, and comparative jury studies.

Before his death in 1916, at age 53, Hugo Muensterberg conducted research on several of the topics discussed in this section, namely, assessment, detection of deception, bystander and witness accuracy, and jury persuasion. At that time, concern was widespread that he was pushing the state of scientific knowledge beyond its limits. Each of the forensic topics explored by Muensterberg has been the subject of extensive inquiry and refinement by psychologists in the ensuing century. As noted previously, concerns that forensic psychologists must work within the limits of scientific knowledge, in particular by scrutinizing the external and ecological validity of their findings, have remained constant. Many forensic topics that Muensterberg presaged through his interest in environmental influences on individual differences have since been added to the repertoire of psychology and law studies, for instance, interventions to treat violent and sex offenders and the identification of individual and environmental factors that lead to offending behavior. Contemporary events have stimulated advances on topics such as the psychology of law enforcement and of terrorism. Newly emerging topics, such as restorative justice and therapeutic jurisprudence, with more emphasis on contextual factors and community involvement derive

from closer interdisciplinary collaboration between scholars trained in law and psychology than was typical in the field in the past. Research, practice, and policy development borne of collaborative endeavours by those trained in psychology and law may meet with less resistance than did Muensterberg's early work in this field.

See Also the Following Articles

Child Custody ■ Child Testimony ■ Eyewitness Identification ■ Eyewitness Testimony ■ Forensic Mental Health Assessment ■ Interrogation and Interviewing ■ Jury Decision Making ■ Legal Competency

Further Reading

Bausermann, R. (2002). Child adjustment in joint custody versus sole custody arrangements: A meta-analytic review. *Journal of Family Psychology*, *16*, 91–102.

Bornstein, B. (1999). The ecological validity of jury simulations: Is the jury still out? *Law and Human Behavior*, *23*, 75–91.

Costanzo, M. (2004). *Psychology applied to law*. Belmont, CA: Wadsworth/Thompson.

Devine, D. J., Clayton, L. D., Dunford, B. B., Seying, R., & Pryce, J. (2001). Jury decision making: 45 years of empirical research on deliberating groups. *Psychology, Public Policy, and Law*, *7*, 622–727.

Finkel, N. J. (1995). *Commonsense justice: Jurors' notions of the law*. Cambridge, MA: Harvard University Press.

Fisher, R. P. (1995). Interviewing victims and witnesses of crime. *Psychology, Public Policy and Law*, *4*, 732–764.

Foote, W. E., & Goodman-Delahunty, J. (2004). *Evaluating sexual harassment: Psychological, social, and legal considerations in forensic examinations*. Washington, DC: American Psychological Association.

Garven, S., Wood, J. M., Malpass, R. S., & Shaw, J. S., III (1998). More than suggestion: The effect of interviewing techniques from the McMartin Preschool case. *Journal of Applied Psychology*, *83*, 347–359.

Grisso, T. (2002). *Evaluating competencies* (2nd ed.). New York: Kluwer Academic/Plenum Press.

Gudjonsson, G. (1999). The making of a serial false confessor: the confessions of Henry Lee Lucas. *The Journal of Forensic Psychiatry*, *10*, 416–426.

Heilbrun, K. (2001). *Principles of forensic mental health assessment*. New York: Kluwer Academic/Plenum Press.

Heilbrun, K., Marczyk, G., & DeMatteo, D. (2002). *Forensic mental health assessment: A casebook*. New York: Oxford University Press.

Hetherington, E. M., Bridges, M., & Insabella, G. M. (1998). What matters? What does not/Five perspectives on the associations between martial transitions and children's adjustment. *American Psychologist*, *53*, 167–184.

Home Office (2002). *Achieving best evidence in criminal proceedings: Guidance for vulnerable or intimidated witnesses, including children*. London: Home Office Communication Directorate.

Krauss, D. A., & Sales, B. D. (2000). Legal standards, expertise, and experts in the resolution of contested child custody cases. *Psychology, Public Policy, and Law*, *6*, 843–879.

Landy, F. J. (1992). Hugo Muensterberg: Victim or visionary? *Journal of Applied Psychology*, *77*, 787–802.

Melton, G., Petrila, J., Poythress, N., & Slobogin, C. (1997). *Psychological evaluations for the courts: A handbook for mental health professionals and attorneys* (2nd ed.) New York: Guilford Press.

Monahan, J., Steadman, H., Silver, E., Appelbaum, P., Robbins, P., Mulvey, E., Roth, L., Grisso, T., & Banks, S. (2001). *Rethinking risk assessment: The MacArthur study of mental disorder and violence*. New York: Oxford University Press.

Muensterberg, H. (1907). *On the witness stand*. New York: Doubleday, Page & Company.

Neitzel, M. T., McCarthy, D., & Kern, M. J. (1999). Juries: The current state of the empirical literature. In R. Roesch, S. D. Hart, & J. R. P. Ogloff (Eds.), *Psychology and law: The state of the discipline* (pp. 23–54). New York: Kluwer/Plenum.

Nicholson, A. (1999). Forensic assessment. In R. Roesch, S. D. Hart, & J. R. P. Ogloff (Eds.), *Psychology and law: The state of the discipline* (pp. 121–173). New York: Kluwer/Plenum.

Perlin, M. L. (2000). *The hidden prejudice: Mental disability on trial*. Washington, DC: American Psychological Association.

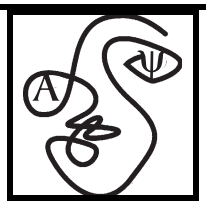
Poythress, N. G., Bonnie, R. B., Monahan, J., Otto, R. K., & Hoge, S. K. (2002). *Adjudicative competence: The MacArthur studies*. New York: Kluwer/Plenum.

Rogers, R. (1997). *The clinical assessment of malingering and deception*. New York: Guilford.

Silver, E., Cirincione, C., & Steadman, H. (1994). Demythologizing inaccurate perceptions of the insanity defense. *Law and Human Behavior*, *18*, 63–70.

Vidmar, N. (2004). Experimental simulations and tort reform: Avoidance, error and overreaching in Sunstein et al.'s. *Punitive Damages* (2002). *Emory Law Journal,* in press.

Zapf, P., & Roesch, R. (1997). Assessing fitness to stand trial: A comparison of institution-based evaluations and a brief screening interview. *Canadian Journal of Community Mental Health*, *16*, 53–66.



Psychometric Tests

Peter F. Merenda
University of Rhode Island, Kingston, Rhode Island, USA

1. Psychometric Tests Defined
2. Early History of Psychometrics
3. Psychometric Theory and Principles
4. Requirement of Technical Manual for All Psychometric Tests
5. Consequences of Using Tests That Do Not Comply with Testing Standards and Ethical Principles
6. Conclusions
Further Reading

GLOSSARY

measurement principles and scales Psychological scaling methods for scoring tests.
objective scoring Procedures that yield the same exact score regardless of who or what conducts the scoring of tests.
projective techniques Measurement instruments and devices in which the stimulus items are unstructured and/or ambiguous.
psychological testing A procedure that involves the use of tests to measure psychological constructs of individual persons.
psychometrics The formal study and practice of test construction principles and evaluation of numerical test results and the interpretation of protocols.
psychophysics The study and determination of sensory thresholds.
subjective scoring Procedures in the scoring of tests that involve intuition, bias, and other subjective judgments that yield different test scores depending on who or what does the scoring.
technical manual Booklet accompanying tests that provides technical information relating to sound psychometric properties of the tests.
testing standards Standards prepared periodically by professional associations with expertise in psychological testing and made available to test constructors and users by the American Psychological Association.

1. PSYCHOMETRIC TESTS DEFINED

The term "psychometric" is defined in *Webster's* dictionary as "the measurement of the duration, force, interrelations, or other aspects of mental processes, as by psychological tests." However, not all psychological tests are psychometric tests. Cronbach, an early prominent psychometrician and author of books and journal articles in the field of psychological testing, identified psychometric tests as those that obtain numerical estimates of human behavior and performance, in accordance with Thorndike's dictum that "if a thing exists, it exists in some amount." Cronbach went on to distinguish psychometric testing from impressionistic testing. Within Cronbach's framework and classification of tests in the former category are tests of general intelligence, ability achievement, performance, personality, and the like. In the latter category are the projective devices and techniques as well as rating scales and interview measures based on subjective judgments. An even earlier distinction between these classifications of tests was formally made by the American Psychological Association (APA) in its first publication of *Ethical Standards of Psychologists* and citing incidents of reported ethical problems of psychologists. In Section 5.4,

©2004 Elsevier Inc.
All rights reserved.

"Publishing and Using Psychological Tests and Diagnostic Aids," it was inferred that the former are psychometric tests and the latter are not. (Incidentally, it was in this section that the three-tier (i.e., A, B, C) hierarchy in classifying tests and diagnostic aids was established and first published.)

1.1. Characteristics of Psychometric Tests

Paramount among the features of psychometric tests is objectivity of scores obtained through standardized testing procedures. This means that whatever or whoever produces the scores, they are always the same; subjective judgment never enters into the process of scoring. The numerical scores of psychometric tests yield continuous or discrete frequency distributions that may legitimately be analyzed statistically. Furthermore, for a test, a scale, or an inventory to qualify as being a psychometric one, there must be evidence that the instrument possesses sound psychometric properties. Finally, the instrument should be shown to measure psychological constructs, particularly those consistent with well-known psychological theory.

1.2. Characteristics of Nonpsychometric Tests

A distinguishing feature, particularly of projective techniques, is the unstructured tasks that are presented to respondents and the ambiguous stimulus items. The principal contrasting feature between psychometric and nonpsychometric tests is the objectivity versus subjectivity of scoring. In the vernacular of psychometricians and measurement specialists, the scoring of nonpsychometric tests involves, in the main, what are referred to as "subjective–intuitive procedures." However, in spite of these differences, instruments in this category are required by national testing standards to be in compliance with those published by national professional societies. Table I provides examples of both psychometric and nonpsychometric tests.

TABLE I
Examples of Psychometric and Nonpsychometric Tests

Psychometric tests	*Nonpsychometric tests*
California Personality Inventory (CPI)	Children's Apperception Test (CAT)
Differential Aptitude Tests (DAT)	Machover Draw-A-Person Test (D-A-P)
Scholastic Assessment Test (SAT)	Rorshach Inkblots (RI)[a]
State–Trait Anxiety Inventory (STAI)	Rosenzweig Picture Frustration Study (RPFS)
Wechsler Intelligence Scale For Children (WISC)	Thematic Apperception Test (TAT)

[a]To be distinguished from the Holtzman Inkblot Technique, for which a genuine attempt has been made by the author, Wayne Holtzman, to meet standards of psychometric tests.

2. EARLY HISTORY OF PSYCHOMETRICS

In the social sciences, particularly education, psychology, and sociology, psychometrics as an academic field of study is of relatively recent origin. It is barely 50 years old. The first formal programs to train psychometricians in graduate schools in the United States were established shortly after the end of World War II. They were the programs at the University of Chicago (headed by Palmer O. Johnson) and at Harvard University (headed by Phillip J. Rulon). However, the underlying roots can be traced to the 19th century, even prior to the establishment of psychology as a discipline by Wundt in 1879.

2.1. Derivation From Psychophysics

Historically, psychometrics can be viewed as an outgrowth of psychophysics during the latter part of the 19th century. In Germany, Fechner and his colleague Weber, who in 1834 had postulated the famous Weber's Law of just noticeable differences, were the founders of psychophysics. Both were early physiologists and were credited with founding psychophysiology. It was these original works by Fechner and Weber that led to the postulation of the measurement model in the United States by Stevens in 1946. But because Fechner's works laid the early foundation to the basic principles of measurement, Fechner is often referred to as the first psychometrician.

2.2. The Measurement Model

At Harvard University, Stevens was the psychologist who provided the social sciences with a taxonomy of measurement scales that form the base for the scoring

of psychometric tests. Stevens had been greatly influenced by the earlier work of Fechner and Weber with physical scales. Besides Stevens, other psychologists in the United States, such as Thurstone and Likert, had already begun to extend Fechner's work to measure psychological constructs. Later, fundamental advances to furthering work in psychometrics beyond psychophysical theory and methods were the measurement models based on true score theory and item response theory.

3. PSYCHOMETRIC THEORY AND PRINCIPLES

Commensurate with and preceding the existence of formal academic graduate courses and programs in psychometrics, as well as the publication and distribution of national testing standards and ethical principles governing the rise of tests operationally and in research, a number of textbooks explaining psychometric theory began to be published. Prominent among these early texts were those of Guilford in 1936 and 1954, Gulliksen in 1950, Lord and Novick in 1968, and Nunnally in 1967. All of these books are still in print and available to students, professors, and researchers. The Lord and Novick book was supplemented by that of Wainer and Braun in 1988. In 1978, Nunnally revised the first edition of *Psychometric Theory,* and several years after his untimely death, a third edition by Nunnally and Bernstein appeared on the scene in 1994.

3.1. Test and Testing Standards

The first test and testing standards were published by APA in 1954, that is, 1 year after its publication of the *Ethical Standards of Psychologists*. These testing standards were the *Technical Recommendations for Psychological Tests and Diagnostic Techniques*. They were published as a supplement to the *Psychological Bulletin*. Subsequently, the American Educational Research Association (AERA) and the National Council on Measurement in Education (NCME) collaborated with APA in producing the *Standards for Educational and Psychological Tests* in 1966, 1974, 1985, and 1999. The documents are published and distributed by APA. They set forth the principles of sound psychometric properties to be met by all psychometric tests as well as all nonpsychometric tests.

3.2. Ethical Principles Governing Use of Tests and Testing Practices

Following the first publication of the *Ethical Standards of Psychologists* in 1953, APA has periodically revised and updated these principles. In each case, a section has been included that relates to the proper use of tests and testing practices. On June 1, 2003, an expanded *Ethical Principles of Psychologists and Code of Ethics* took effect.

3.3. Principal Psychometric Properties of Tests

The current testing *Standards* publication is the most comprehensive single documentation of tests and testing practices ever published and circulated by APA. There are 15 sections of standards that cover 194 pages (including the glossary and index). Foremost among the standards with which all psychometric (and other measurement) instruments must comply are Standards 1 to 6:

1. Validity
2. Reliability and errors of measurement
3. Test development and revision
4. Scales, norms, and score comparability
5. Test administration, scoring, and reporting
6. Supporting documentation for tests

Adhering to the principles espoused by the *Standards* for testing and mandated by the *Ethical Principles* leads to the assurance that measurement instruments and the scores they yield possess the required psychometric properties. Only in this way can it be established that the assessments measure the psychological constructs that are purported to be measured by the instruments. Equally important is evidence that adherence leads to a determination that the scores produced are internally consistent and stable over time (i.e., reliability) and that the purposes for which the tests were developed are being fulfilled. Compliance also ensures that the properties of the tests demonstrate evidence of and support for the meaningfulness and usefulness of the interpretation of the scores yielded by the tests and are consistent with psychological theory (i.e., validity).

4. REQUIREMENT OF TECHNICAL MANUAL FOR ALL PSYCHOMETRIC TESTS

In Section 6 of the *Standards,* authors and test publishers are advised to provide test users with empirical

documentation that the tests in question possess the required sound psychometric properties. Standard 6.1 states, "Test documents (e.g., test manuals, technical manuals, user's guides, and supplemental materials) should be made available to prospective test users and other qualified persons at the time a test is published or released for use." Until recently, many psychometricians and measurement specialists who were requested to review and critique tests for the *Mental Measurement Yearbooks,* including the current author, would decline to review tests that were not accompanied by technical manuals and other supporting documentation. The 2001 edition of the yearbook contains an appendix titled "Tests Lacking Sufficient Technical Documentation for Review" listing 94 such tests that had been submitted for review to the Buros Institute of Mental Measurements.

5. CONSEQUENCES OF USING TESTS THAT DO NOT COMPLY WITH TESTING STANDARDS AND ETHICAL PRINCIPLES

Prior to the publication by APA of the first ethical codes of psychologists in 1953 and the first testing standards in 1954, there were no national regulations in the United States governing the use of tests, whether psychometric or otherwise. However, the current author began to testify in the federal courts as an expert witness (a psychometrician) to support the sound psychometric properties of tests in 1952. Over the subsequent 50 years, he continued to testify in favor of the proper use of psychometric tests and in opposition to the misuse of tests. Much of the misuse has involved violations of the testing standards and the ethical codes of psychologists. These violations have included failure to accompany tests with technical manuals and further documentation, improper application of tests, improper interpretation of test scores, and "stealing" of materials and systems copyrighted by legitimate test authors.

The consequences of violations of the *Standards* and *Ethical Principles* may fall into moral or legal categories. Moral violations may be those on the part of psychologists who are APA members and/or American Psychological Society members. Violations in this category usually result in dismissal of membership in either or both professional associations as well as loss of license for psychologists in private practice. (The same is true for members of other professional associations and societies.) The consequences in the legal category are much more severe and costly, especially in court decisions in class-action suits. As an example of the latter, a summary of a class-action suit in which the current author was the principal expert witness (as a psychometrician) for the plaintiffs is presented next.

A 1991 case, *Soroka, et al. v. Dayton–Hudson Corporation,* was settled by an out-of-court agreement between the parties. The consent decree was ordered by the trial court judge and was signed by both parties on September 23, 1993. The general terms of the agreement provided for payment of $1.3 million to an estimated 2500 job applicants who were administered the test at issue between 1988 and 1991, payment of an additional $60,000 for four plaintiffs legal fees, and payment of an additional $35,000 for services to the plaintiffs' legal staff in administering the claims procedure and the settlement funds.

The case was filed by applicants for the position of store security officer at Target Stores in California. Target Stores is owned and operated by Dayton–Hudson, with headquarters in Minneapolis, Minnesota. The applicants claimed that they were required to take the 704-item Rodgers Condensed California Personality Inventory–Minnesota Multiphasic Personality Inventory (CPI–MMPI) and alleged that the test was not psychometrically sound, that it violated California's fair employment laws, and that it constituted an unlawful invasion of privacy protected by the constitution of the State of California.

Four expert witnesses for the plaintiffs (three psychometricians and one industrial–organizational psychologist) were successful in convincing the court of the merits of the complaints and charges filed against Dayton–Hudson. It was the experts' respective testimonies, based primarily on alleged violations by the defendant, that led Dayton–Hudson to seek the out-of-court agreement. On October 25, 1991, the court granted the plaintiffs a preliminary injunction on further testing in concluding that the plaintiffs had in fact convinced the court of the merits of the claims that had been filed. The defendant had considered appealing the ruling to the California Supreme Court but then decided against following through on the appeal.

6. CONCLUSIONS

Psychometrics has enjoyed a history about as long as that of psychology itself. It has been viewed as an outgrowth of the earlier psychophysics. Many

psychometricians and measurement specialists recognize Fechner, the early German psychophysiologist, as the Father of Psychometrics.

Psychological tests are classified into two major categories: psychometric and nonpsychometric. The principal distinctions between the two rest mainly with (a) the nature of the scoring of test items and (b) the nature of the items themselves. To qualify as a psychometric test, the scoring procedures must be purely objective, whereas in the scoring procedures for nonpsychometric tests, subjective judgment on the part of the scorers enters into the process of scoring the items. Relating to the items themselves, those for psychometric tests are structured, whereas those for nonpsychometric tests are unstructured and generally ambiguous. However, both types of psychological tests are required by national regulations to comply with the *Ethical Standards of Psychologists* and with the *Ethical Principles of Psychologists and Code of Conduct*. The publication and dissemination of these testing standards and ethical principles and codes for test constructors and test users dates back more than 50 years. Violation of these regulatory standards and principles may result in dire consequences to the violators so as to protect the test takers.

See Also the Following Articles

Assessment and Evaluation, Overview ■ Intelligence Assessment ■ Personality Assessment ■ Psychological Assessment, Standards and Guidelines for ■ Psychophysiological Assessment

Further Reading

American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. (1966). *Standards for educational and psychological tests and manuals*. Washington, DC: Authors.

American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. (1974). *Standards for educational and psychological tests*. Washington, DC: Authors.

American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. (1985). *Standards for educational and psychological testing*. Washington, DC: Authors.

American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. (1999). *Standards for educational and psychological measurement*. Washington, DC: Authors.

American Psychological Association. (1953). *Ethical standards of psychologists*. Washington, DC: Author.

American Psychological Association. (1954). *Technical recommendations for psychological tests and diagnostic techniques*. Washington, DC: Author.

American Psychological Association. (1992). *Ethical principles of psychologists and code of conduct*. Washington, DC: Author.

Buros Institute of Mental Measurement. (2001). *Mental Measurement Yearbook* (14th ed.). Lincoln: University of Nebraska Press.

Camara, W. J., & Merenda, P. F. (2000). Using personality tests in employment screening: Issues raised in *Soroka v. Dayton–Hudson Corporation*. *Psychology, Public Policy, and Law*. 6, 1164–1186.

Cronbach, L. J. (1960). *Essentials of psychological testing* (2nd ed.). New York: Harper & Row.

Fechner, G. (1860). *Elements de Psycholigsik*. Leipzig, Germany: Bretikopf & Hartel.

Guilford, J. P. (1936). *Psychometric methods*. New York: McGraw–Hill.

Guilford, J. P. (1954). *Psychometric methods* (2nd ed.). New York: McGraw–Hill.

Gulliksen, H. (1950). *Theory of mental tests*. New York: McGraw–Hill.

Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology*, *140*, 44–53.

Loevinger, J. (1947). A systematic approach to the construction and evaluation of tests of ability. *Psychological Monographs*, *61*(No. 4).

Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Reading, MA: Addison–Wesley.

Merenda, P. F. (1990). Present and future issues in psychological testing in the United States. *Psychological Assessment*, *6*, 3–31.

Merenda, P. F. (1993). An anticipated landmark decision involving issues related to psychological assessment of applicants for positions in business and industry. *Psychological Reports*, *72*, 102.

Merenda, P. F. (1994). The settlement of the "Target" case and its aftermath. *Psychological Reports*, *75*, 1485–1486.

Merenda, P. F. (1995). Substantive issues in the *Soroka v. Dayton–Hudson* case. *Psychological Reports*, *77*, 595–606.

Merenda, P. F. (2003). Measurements in the future: Beyond the 20th century. *Psychological Reports*, *92*, 209–217.

Nunnally, J. (1967). *Psychometric theory*. New York: McGraw–Hill.

Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York: McGraw–Hill.

Nunnally, J. C., & Bernstein, L. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw–Hill.

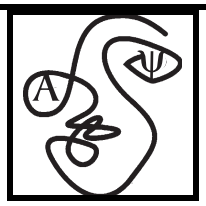
Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen, Denmark: Paedgogiske Institute.

Sereci, S. G., Wainer, H., & Braun, H. (1998). *Psychometrics*. Princeton, NJ: Educational Testing Service.

Stevens, S. S. (1946). On the theory of scales of measurement. *Science*, *103*, 667–680.

Thurstone, L. L. (1927). A law of comparative judgment. *Psychological Review*, *34*, 273–286.

Wainer, H., & Braun, H. J. (Eds.). (1988). *Test validity*. Mahwah, NJ: Lawrence Erlbaum.



Psychoneuroimmunology

Lise Solberg Nes and Suzanne C. Segerstrom

University of Kentucky, Lexington, Kentucky, USA

1. About the Immune System
2. Pathways between the Brain and the Immune System
3. Stress, Hormones, and Immunity
4. Psychological and Behavioral Interventions
5. PNI and Interventions for Chronic Disease
6. Conclusion

Further Reading

GLOSSARY

cytokines Proteins created by cells that influence the behavior of other cells; they often act as messenger molecules that enable communication between cells, particularly immune cells.

endocrinology The study of internal secretions of the body, particularly hormonal secretions, and activities involved in such secretions.

granulocytes White blood cells containing granules that themselves contain substances that can be used to combat invaders (e.g., oxygen radicals).

hormone A substance secreted by the body that affects distant target cells.

immune system A system of organs and cells that protects the body from intruding organisms and, in some cases, malfunctioning cells.

interventions Attempts to modify behavior, cognition, or emotions so as to improve psychological and physical well-being.

leukocytes White blood cells, including granulocytes, monocytes, and lymphocytes.

lymphocytes White blood cells that have specific abilities to kill particular invaders (e.g., a given cell recognizes and responds to only one stimulus) through cytotoxic action or antibody production; lymphocytes include B and T cells and natural killer (NK) cells, which are a type of nonspecific lymphocyte.

monocytes Mononuclear cells that can migrate into tissue (where they are called macrophages), induce inflammation, and phagocytose (i.e., engulf or eat) invaders.

neuroendocrine Relationships between the nervous and endocrine systems as well as cells that release hormones into the blood in response to a neural stimulus.

psychoneuroimmunology The study of interactions among the mind, the nervous system, and the immune system.

psychosocial Having to do with an individual's psychological or social state.

Psychoneuroimmunology (PNI) is the study of interactions among the mind, the nervous system, and the immune system. The mind, or psyche, involves thoughts, emotions, experiences, and ideas; the nervous system involves the brain, the spinal cord, and nerves throughout the body; and the immune system consists of organs and cells that defend the body against invaders. The mind and the immune system communicate through the peripheral nervous system, hormones, and cytokines. This communication allows the immune system to be responsive to psychosocial factors and allows the immune system to signal the brain (e.g., about infection). The field of PNI is interdisciplinary and

©2004 Elsevier Inc.
All rights reserved.

connects knowledge from endocrinology, immunology, psychology, neurology, and other areas. There is increasing interest in the applied aspects of PNI, specifically whether psychosocial interventions can affect diseases such as human immunodeficiency virus (HIV) and cancer.

1. ABOUT THE IMMUNE SYSTEM

The main function of the immune system is to protect the body from infections. When foreign substances or pathogens (e.g., bacteria, viruses, parasites, fungi) attack the body, the immune system detects and destroys these invaders. In addition, the immune system can repair and heal tissue damage from injury as well as destroy and eliminate otherwise normal cells that have been altered by infections or, in some cases, malignancies.

The immune system consists of white blood cells that develop in the bone marrow and the thymus; congregate in the spleen, the lymph nodes, the lungs, the gastrointestinal (GI) tract, and the skin; and circulate through the blood and lymph. These cells include granulocytes and macrophages, which act nonspecifically in processes such as inflammation. They also include lymphocytes, such as B, T, and natural killer (NK) cells, which act specifically on targeted invaders to kill infected cells (i.e., cytotoxicity) and produce proteins that attack invaders (i.e., antibodies). A critical element of the immune response is the production of molecules that amplify and direct the immune response (i.e., cytokines). Different cytokines promote inflammation, cytotoxicity, antibody production, or some combination thereof.

2. PATHWAYS BETWEEN THE BRAIN AND THE IMMUNE SYSTEM

Historically, the immune system was thought to be an independent system. However, during the 1960s and 1970s, evidence that the immune system could be influenced by psychological phenomena began to accumulate. In 1964, Solomon and Moos proposed that the development of rheumatoid arthritis was related to personality factors, advancing the possibility of psychoneuroimmunological phenomena. In 1975, Ader and Cohen demonstrated that conditioning could affect the immune response by pairing taste with an immunosuppressive drug. Animals that were conditioned in this way showed evidence of immunosuppression when the taste was later presented alone. Evidence of anatomical pathways between the brain and the immune system began with the discovery of sympathetic nervous system innervation of immune organs and has accumulated rapidly to include numerous neurological, hormonal, and cytokine pathways.

2.1. Autonomic Nervous System

Bidirectional neurological pathways exist between the central nervous system (CNS) and the immune system. The most direct pathway from the CNS to the immune system is through the autonomic nervous system (ANS) as sympathetic nerve fibers innervate lymphoid organs such as the thymus, the bone marrow, and the spleen. The influence on the immune system depends on the specific cells and receptors involved and can be inhibitory or stimulatory. Recent evidence suggests that the immune system can also signal the brain through the ANS, specifically the subdiaphragmmatic vagus nerve.

2.2. Endocrine System

The endocrine system consists of a variety of glands that secrete hormones. Hormones are chemical substances that are released into the bloodstream, stimulate receptors in target cells, and alter the behavior of those cells. Ample evidence exists that neuroendocrine substances influence immune cells. For example, immune cells react to many hormones, including cortisol, substance P, vasoactive intestinal peptide, prolactin, and growth hormone. Many of the cells involved in the immune response carry surface receptors for a number of hormones, allowing them to respond to signals from the brain as well as from other organ systems.

2.3. Immune-to-Brain Communication

Cytokines, particularly proinflammatory cytokines, can affect both neuroendocrine and brain function in an endocrine fashion. When an immune response occurs, cytokines signal a state of disease to the CNS. Proinflammatory cytokines raise body temperature (fever), and this retards the growth of some bacteria but does so at a cost: Fever increases the host's metabolic rate by 13% for each 1 degree Celsius. As a result, behavior changes to preserve metabolic resources for fighting the infection. "Sickness behavior," which

accompanies fever, includes feeling tired and weak, loss of interest in surroundings, loss of appetite, and loss of concentration. These behavioral changes are consistent with centrally mediated change in priorities away from activity and pursuit of goals to conservation of energy.

3. STRESS, HORMONES, AND IMMUNITY

The idea that stressor exposure might influence the immune response was an early topic of interest in psychoneuroimmunology (PNI), and evidence clearly indicates that psychological and physical stressor exposure can affect the immune system. Both the type of stress and the person involved may affect the influence of stress on the immune system. Stress can be physical or psychological, controllable or uncontrollable, and acute or chronic. Furthermore, different people may handle stressor exposure in different ways.

When a situation is perceived as threatening, the sympathetic nervous system responds by releasing epinephrine and norepinephrine. Consequences of this response include increased heart rate, increased blood pressure, and peripheral vasoconstriction. Stressor exposure also influences the endocrine response, changing levels of hormones such as cortisol, a steroid hormone produced by the hypothalamic–pituitary–adrenal (HPA) axis. These responses can, in turn, affect the immune system.

There are more than 300 studies that have linked stress and stressors to immune parameters. In a meta-analysis of these studies, both acute stressors (lasting minutes) and chronic stressors (lasting days to years) were reliably associated with changes in the immune system. The effects of stressors and stress hormones depend critically on characteristics of the stressor, particularly time course. In general, the changes associated with acute stressors could be adaptive as part of the "fight or flight" response; acute stressors (lasting minutes) are associated with large increases of nonspecific defenders (e.g., neutrophils, NK cells) in the bloodstream, release of antibody into saliva, and increased production of cytokines that stimulate nonspecific immunity. Mechanistic investigations have shown that many of these changes are due to sympathetic signaling through β2-adrenergic receptors on lymphocytes.

However, longer term stressors had less potentially beneficial effects. The shortest "chronic" stressors, such as academic examinations, were associated with decreases in the function of cytotoxic lymphocytes (e.g., NK cells, T cells), but antibody-producing functions were generally unchanged. Chronic stressors that were long-lasting, such as having a spouse with dementia, decreased both cytotoxic and antibody-producing functions of immune cells.

4. PSYCHOLOGICAL AND BEHAVIORAL INTERVENTIONS

PNI is mainly a basic science. However, the number of studies investigating the impact of psychosocial interventions on the immune system is increasing. In particular, the hypothesis and evidence that psychosocial stress can affect the immune system has generated the secondary hypothesis that psychosocial interventions to alleviate stress could have the opposite effect on the immune system. Such interventions commonly involve stress management, relaxation, and hypnosis. A recent meta-analysis found that there were few effects of interventions on the immune system but also that most stress management and relaxation interventions failed to enroll stressed participants. In those that did, there was some evidence that the interventions could alleviate stressor-related immune change.

4.1. Stress Management

In stress management interventions, participants are taught to recognize situations they experience as stressful and to identify the reactions they have in these situations. The participants are then trained to avoid unnecessary stressor exposure and to cope better with stressors. Stress management interventions have mostly been targeted to patients with chronic medical illnesses such as cancer and HIV. Such interventions usually include information and education about the illness, access to psychological support, help in changing cognition about the disease, and training to improve coping skills. Studies involving stress management interventions have yielded mixed results, but some such interventions have been found to reduce negative emotions and have been associated with positive changes in immune function, including increases in CD4+ T lymphocyte numbers, NK cell numbers, and NK cell activity.

4.2. Disclosure and Health

Some data indicate that undisclosed traumatic cognitions or emotions might lead to serious physiological consequences. In intervention studies on this topic, participants are asked to write or talk about stressful or traumatic situations that they have experienced. The cognitive processing and emotional expression during this activity are thought to lead to better regulation of physiological systems. Research in this area has found associations between disclosure and reduced negative affect, reduced activity in the sympathetic nervous system, and reduced health center visits. Immunological consequences have included better response to vaccination and better immune control of latent viruses.

4.3. Relaxation

A relaxed state can be induced in several ways, including progressive muscle relaxation, relaxation response, biofeedback, imagery, hypnosis, and meditation. In general, all of these forms of relaxation have been related to decreased negative mood, decreased hormone levels, and decreased activity in the sympathetic nervous system. However, the results of relaxation techniques are inconsistent with regard to immune responses. Relaxation exercises are associated with increased levels of antibody in saliva and may increase levels of CD4 helper T cells in the blood as well as NK cell activity.

4.4. Hypnosis

Hypnosis involves the induction of an altered state of consciousness, including focused attention and high suggestibility. In research studies, the application of hypnosis to influence the immune system is often through hypnotic suggestion to affect an allergic response. The participant first has an allergen injected under the skin. A hypnotic state is then induced, and the hypnotized participant is instructed to either enhance or suppress his or her response to the allergen. In most hypnosis interventions, the direction of effect is specified. Following such interventions, hypnosis has successfully suppressed, but not enhanced, the immune response. Studies have also shown that hypnotic suggestions can lead to changes in the ANS such as alterations in heart rate and blood pressure. In medical contexts, hypnosis has been used to manage chronic pain and to enhance adjustment to physical conditions and illnesses. There are great differences in people's susceptibility to hypnosis, and highly hypnotizable individuals have the strongest immune changes.

5. PNI AND INTERVENTIONS FOR CHRONIC DISEASE

5.1. Interventions in HIV

The clinical course of HIV infection varies greatly, and although some people can live with HIV infection for years without developing acquired immune deficiency syndrome (AIDS), others progress more quickly after being infected. A number of factors, such as viral strain, genetic immune factors, presence of other infections, nutrition, and drug abuse, are thought to influence progression. In addition, a host of psychosocial stressors related to the diagnosis and stigma of the illness may contribute to differences. A person infected with HIV may face social, medical, legal, professional, and mental problems on diagnosis, and coping with these stressors may make a difference in how the illness progresses.

HIV infection develops into AIDS when the immune system is too weak to fight infections. HIV infection has been associated with physical and psychological stressor exposure and distress, and because stressor exposure and distress have been associated with suppression of immune functions, the progression of HIV infection may be influenced by interventions that can affect responses to stress. In intervention studies, HIV-infected patients in cognitive–behavioral stress management programs had decreased distress levels and higher NK cell and CD4+ levels than did HIV-infected controls. HIV-infected patients in cognitive–behavioral stress management programs also had decreases in depression and anxiety, and both cognitive–behavioral training and relaxation training appear to have positive impact on immune function among patients. However, findings from other studies have yielded contradictory results, and although relationships among behavior, psychosocial stressors, immune status, and HIV progression apparently exist, these relationships are not completely understood.

5.2. Interventions in Cancer

PNI may have implications for cancer. One possibility is that psychosocial factors influence the onset and especially the advancement of cancer through psychological impact on the immune system. This is the most

traditional view, and psychological influences on the immune system are often assumed to be the connection between psychosocial factors and cancer. However, controversy surrounds this topic because few studies have been able to detect clear associations between psychological effects on the immune function and cancer outcomes. Some of this confusion may arise from the heterogeneity among various types of cancer that have different molecular mechanisms and different immunological vulnerabilities.

PNI may also be of importance to oncology in that psychological influences on the immune system may affect the risk of infectious disease. Infectious disease is a major source of concern in oncology because the immune suppression often seen in patients due to cancer treatments and psychological distress may render patients' immune systems incapable of fighting off intruding viruses and bacteria, leaving patients more susceptible to infectious disease. However, there is not yet much research in this area.

Associations have been found between support group interventions and longevity in people with cancer. However, other studies have found no association. Some interventions have demonstrated an immune benefit for people with cancer. In cognitive–behavioral stress management interventions, breast cancer and malignant melanoma patients received health education, stress management training, problem-solving skills training, and social support. Participants had decreased depression and anxiety as well as increased immune parameters. However, such effects remain to be replicated, and the consequences of these changes for patient health remain unclear.

6. CONCLUSION

Clear connections exist between the nervous and neuroendocrine systems and the immune system, and compelling evidence points to the influence of psychosocial factors on immune function as well as vice versa. Interventions to reduce potentially deleterious effects of stress and depression on the immune system are in their infancy, but there are encouraging examples of how applied PNI may contribute to better immune function and, potentially, to better health.

See Also the Following Articles

Health and Culture ■ Health Psychology, Cross-Cultural ■ Hypnosis ■ Stress

Further Reading

Ader, R., & Cohen, N. (1975). Behaviorally conditioned immunosuppression. *Psychosomatic Medicine*, *37*, 333–340.

Ader, R., Felton, L., & Cohen, N. (2001). *Psychoneuroimmunology* (3rd ed.). San Diego: Academic Press.

Fawzy, F. I., Fawzy, N. W., & Hyun, C. S. (1994). Short-term psychiatric intervention for patients with malignant melanoma: Effects on psychological state, coping, and the immune system. In C. E. Lewis, C. O'Sullivan, & J. Barraclough (Eds.), *The psychoimmunology of cancer: Mind and body in the fight for survival* (pp. 292–319). New York: Oxford University Press.

Kiecolt-Glaser, J. K., & Glaser, R. (1992). Psychoneuroimmunology: Can psychological interventions modulate immunity? *Journal of Consulting and Clinical Psychology*, *60*, 569–575.

Maier, S. F., Watkins, L. R., & Fleshner, M. (1994). Psychoneuroimmunology: The interface between behavior, brain, and immunity. *American Psychologist*, *49*, 1004–1017.

Miller, G. E., & Cohen, S. (2001). Psychological interventions and the immune system: A meta-analytic review and critique. *Health Psychology*, *20*, 47–63.

Pennebaker, J. W. (1995). *Emotion, disclosure, and health*. Washington, DC: American Psychological Association.

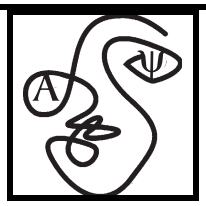
Rabin, B. S. (1999). *Stress, immune function, and health*. New York: Wiley–Liss.

Schedlowski, M., & Tewes, U. (1999). *Psychoneuroimmunology: An interdisciplinary introduction*. New York: Kluwer Academic/Plenum.

Segerstrom, S. C., & Miller, G. E. (in press). Psychological stress and the human immune system: A meta-analytic study of 30 years of inquiry. *Psychological Bulletin*.

Solomon, G. H., & Moos, R. H. (1964). Emotions, immunity, and disease: A speculative theoretical integration. *Archives of General Psychiatry*, *11*, 657–674.

Spiegel, D., Bloom, J. R., Kraemer, H. C., & Gottheil, E. (1989). Effect of psychosocial treatment on survival of patients with metastatic breast cancer. *Lancet*, *2*, 888–891.



Psychophysiological Assessment

Jaime Vila
University of Granada, Granada, Spain

1. Introduction
2. Historical Antecedents
3. Peripheral Psychophysiological Measures
4. Central Psychophysiological Measures
5. The Psychological Context
6. Assessment of Basic Psychological Processes
7. Clinical Applications
Further Reading

GLOSSARY

cardiovascular activity Functional activity of the heart and vessels controlled by the autonomic nervous system.
electrodermal activity Electrical activity of the skin due to activation of the sweat glands by the sympathetic nervous system.
electromyography Electrical activity of the striated muscle controlled by the somatosensory nervous system.
event-related potential Specific change in the electroencephalogram (EEG) due to the presentation of a stimulus or in anticipation of a response; it is recorded after averaging the EEG signals produced by a number of repeated events.
functional magnetic resonance imaging Neuroimaging technique based on the magnetic properties of the spinning protons; the greater the magnetic signal in a brain area, the greater the blood flow and neural activity in that area.
neuropsychological disorder Psychological disorder due to brain lesion.
psychopathological disorder Psychological disorder not due to brain lesion or other organic disorder.
psychophysiological disorder Physiological disorder due to psychological factors.
psychophysiology The scientific study of psychological processes through noninvasive recording of physiological responses.

Psychophysiology is the scientific discipline devoted to the study of the relationships between the physiological and psychological aspects of behavior. Although such relationships have been studied from different perspectives, the psychophysiological perspective emphasizes the use of noninvasive physiological measures—both peripheral and central—to assess the psychological processes underlying human behavior.

1. INTRODUCTION

The scientific study of human behavior becomes enriched when it is done in close relationship with the biological body that sustains behavior. To study the activity of the organisms from this perspective means to examine what people do, think, and feel while simultaneously taking into account the physiological and psychological mechanisms that control human behavior.

The relationships between physiology and psychology have preoccupied philosophers and scientists throughout history. The different debates about the soul–body, spirit–matter, and mind–brain relationships all are variations on a single theme: the connection between the behavior—people's acts, thoughts, and feelings—and its sustaining biological body.

©2004 Elsevier Inc.
All rights reserved.

2. HISTORICAL ANTECEDENTS

Nearly all physiological recording methods are based on the bioelectrical nature of the living tissue activity. Since the discovery of electricity by the end of the 18th century and throughout the electronic and computer revolution of the 20th century, it is not surprising that the history of the psychophysiological approach to assessment evolved in parallel with technological evolution.

The first psychophysiological instruments were based on the galvanometer, that is, a moving coil suspended in the magnetic field of a permanent magnet that rotates when an electric current flows through the coil. A galvanometer-type device used with a writing system mounted on the moving coil is the basis of the traditional polygraphs.

The electronic and computer revolution during the 20th century influenced psychophysiological instruments in three basic aspects: (a) increased precision and reduced size and weight of the amplifiers; (b) improvements in the traditional recording systems using digital computers to represent, store, and analyze the psychophysiological signals, and (c) allowing the discovery of new psychophysiological variables only available after complex computerized systems have been used to detect and extract the biological signals.

Psychophysiological equipment must take into account the physiological recording and the psychological manipulation. Figure 1 illustrates the basic elements of current equipment. On the left, the participant has the sensors placed on the surface of the body while the psychological task is being performed. On the right, in the experimenter's room, are the physiological recording system, the connection box, the analog-to-digital and digital-to-analog card, the computer central processing unit where the card is inserted, the computer monitor where the physiological recording is being displayed, and the slide projector, controlled by the computer, to present the psychological stimuli.

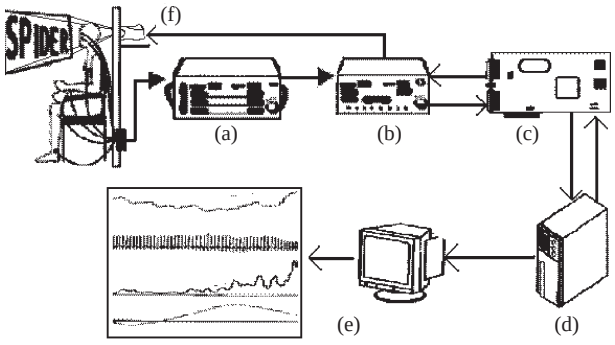


FIGURE 1 Diagram representing the basic elements of a psychophysiological laboratory: (a) the physiological recording system; (b) the connection box; (c) the analog-to-digital and digital-to-analog card; (d) the computer central processing unit where the card is inserted; (e) the computer monitor where the physiological recording is being displayed; and (f) the slide projector, controlled by the computer, to present the psychological stimuli.

3. PERIPHERAL PSYCHOPHYSIOLOGICAL MEASURES

Psychophysiological techniques are classified into peripheral and central techniques. Peripheral techniques include those variables directly controlled by the somatosensory nervous system and the autonomic nervous system (e.g., surface electromyography, eye movements, respiration, cardiovascular activity, electrodermal activity). Central techniques include those variables directly controlled by the central nervous system (e.g., event-related potentials [ERPs], functional magnetic resonance imaging [fMRI], positron emission tomography [PET]). Table I summarizes the various psychophysiological techniques following this classification. This section describes one psychophysiological technique belonging to the peripheral somatosensory system (surface electromyography) and two techniques belonging to the peripheral autonomic nervous system (cardiovascular activity and electrodermal activity).

3.1. Surface Electromyography

The biological basis of electromyography is the electrical activity of the striated muscle, that is, that type of muscle responsible for motor activity and made up of a large number of parallel fibers. Such activity is triggered by the activation of the motoneuron that innervates the fiber. This functional physiological entity, made up of a single motoneuron and the individual muscle fibers innervated by the motoneuron, is called the motor unit. Recording of the electrical activity of the striated muscles from the surface of the skin is referred to as surface electromyography (EMG). This EMG signal represents the ensemble of action potentials produced by the muscle at a given moment. The

TABLE I
Classification of Psychophysiological Techniques

Neurophysiological mechanism	*Psychophysiological techniques*
Somatic nervous system	Surface electromyography
	Direct electromyography
	Integrated electromyography
	Eye movements
	Respiration
Autonomic nervous system	Cardiovascular activity
	Electrocardiogram
	Impedance cardiography
	Photoplethysmography
	Blood pressure recording
	Electrodermal activity
	Pupillary activity
	Gastrointestinal activity
	Others
Central nervous system	Electromagnetic techniques
	Electroencephalogram
	Event-related potentials
	Brain electrical activity mapping
	Magnetoencephalography
	Metabolic techniques
	Positron emision tomography
	Functional magnetic resonance imaging

aggregate signal is characterized by a frequency range between 20 and 1000 hertz and an amplitude range between 2 and 2000 microvolts, displaying a direct relationship between frequency and amplitude; the greater the frequency, the greater the amplitude. In psychophysiological assessment, it is common to transform the direct surface EMG into integrated EMG. The integration procedure implies, first, rectification of the signal (i.e., elimination of the negative mirror part of the signal) and, second, smoothing of the remaining signal by a low-pass filter. This produces a contour following integration, that is, a varying voltage proportional to the envelope of the direct EMG signal.

3.2. Cardiovascular Activity

The cardiovascular system consists of the heart and a network of vessels for distribution, exchange, and return of the blood throughout the body. The cardiovascular system reacts quickly to physical and psychological demands, with such activity being controlled by neural and humoral pathways. The neural pathway to the heart involves the two branches of the autonomic nervous system: the sympathetic and parasympathetic branches. When the sympathetic nerves are activated, the heart rate accelerates and heart contractions become stronger. When the parasympathetic nerves are activated, the heart decelerates. The neural pathway to the vessels involve mainly the sympathetic nerves, and its activation produces either vasoconstriction (in skin and viscera) or vasodilation (in motor muscles). The humoral pathway involves the neural sympathetic activation of the endocrine system, resulting in the secretion by the adrenal glands of various hormones into the bloodstream (e.g., adrenaline, cortisol). The impact of psychological processes on cardiovascular activity is assessed through noninvasive measurements of the various components of cardiovascular activity; heart rate (the number of ventricular contractions per minute), blood pressure (the force of the circulating blood in the arteries, and blood flow (the amount of circulating blood in a particular area). In psychophysiological studies, heart rate is the main index to assess cardiac function. Heart rate is measured in terms of beat-by-beat heart rate values derived from the electrocardiogram (R–R interval).

3.3. Electrodermal Activity

Known since the end of 19th century, electrodermal activity is one of the peripheral measures most frequently used in psychophysiological assessment, with its biological basis being the activity of the eccrine sweat glands. Sweat glands are distributed along the surface of the body with a greater density in the inner parts of the hands and feet. The activities of the glands are controlled by the sympathetic branch of the autonomic nervous system, and when the glands are activated, there is a hydration of the skin accompanied by an electrical signal. The electrical signal can be detected on the surface (called skin potential), and there is an increase in conductivity to the passage of an external current (called skin conductance). This activity responds to thermal and psychological stimuli. The psychological significance of the electrodermal activity is predominant when the temperature is below 25 degrees Celsius. The recording of electrodermal activity can be done using two general procedures. First, the endogenous procedure records the electrical potential produced in the skin without passing any external

current. Second, the exogenous procedure records the change in conductivity to the passage of an external current. Depending on the type of external current (DC or AC) and whether the current maintains a constant voltage or constant intensity, the recording procedure is called skin conductance (DC and constant voltage), skin resistance (DC and constant intensity), skin admittance (AC and constant voltage), or skin impedance (AC and constant intensity).

4. CENTRAL PSYCHOPHYSIOLOGICAL MEASURES

The techniques to measure central nervous system activity are classified into two categories: electromagnetic and metabolic techniques. The electromagnetic techniques are based on the electrical activity of the brain tissue, that is, the neurons' action potentials. The metabolic techniques are based on the differential distribution of blood elements (mainly oxygen and hydrogen) in the brain.

4.1. Electromagnetic Techniques

Electroencephalography (EEG) was the first electromagnetic technique developed. Like all psychophysiological techniques, it requires careful attention to electrode location, choice of electrode reference, grounding, proper filtering and artifact control. A good EEG recording is characterized by changes in voltage as a function of time with two inversely related parameters: frequency and amplitude. The lowest frequency band (0–4 hertz, called the delta range) has the greatest amplitude (20–200 microvolts), and the highest frequency band (36–44 hertz, called the gamma range) has the lowest amplitude (2–20 microvolts). The intermediate frequency bands, called the theta (5–7 hertz), alpha (8–12 hertz), and beta (13–30 hertz) ranges, have a progressively decreasing amplitude. An important technological development based on the EEG occurred during the 1960s with the advent of digital computers. Digital computers allowed the event-related brain activity to be extracted from the background EEG by applying the averaging procedure, that is, repeating a particular event and averaging the digital samples of the EEG that are time-locked to that event. The resulting signal, the ERP, contains a number of positive and negative peaks identified by the letters P and N, respectively, each with a numeric subscript (1/2/3/4 or 100/200/300/400) indicating the order or time in milliseconds of the peak, respectively. Examples of ERPs include the N_{100}, P_{300}, and N_{400}. Once extracted, ERPs are interpreted as manifestations of the specific psychological processes activated by the event.

4.2. Metabolic Techniques

The metabolic techniques are similar to traditional radiological techniques (e.g., X ray, computerized tomography). The neuroimage is elaborated from the recording of radiological signals emitted by moving blood particles in the brain (oxygen and hydrogen) necessary for neuronal activity. The first metabolic neuroimaging techniques developed were the radiotracer techniques. PET is one of them. The radiotracer commonly used in PET is a pharmacological compound labeled with a positron-emitting radioisotope (O_{15}) of short life (approximately 4–10 minutes). Once administered to the participant either intravenously or by inhalation, the radiotracer is distributed through the blood to all parts of the body, including the brain. The number of emitted positrons detected from outside the head reflects the amount of oxygen demand in the various brain areas.

The most promising metabolic technique, as an alternative to radiotracers, is fMRI. The physical foundations of fMRI are the magnetic properties of the spinning protons within the nucleus of certain atoms. These subatomic particles behave like small magnetic needles. When the body is placed within the magnetic field of a potent external magnet, all hydrogen and oxygen protons line up along the axis of the external magnet. The orientated spinning protons rotate, or precess, about the axis of the external magnetic field with a rotating frequency that is within the range of radio frequency waves. Using a transmit–receive coil, the MRI technique then applies a momentary external radio frequency wave perpendicular to the magnet axis and identical to the precessional frequency of the spinning protons to produce the resonance phenomenon, that is, a momentary electromagnetic signal detected by the receive coil and used to create the image.

5. THE PSYCHOLOGICAL CONTEXT

The physiological recording becomes psychophysiological only when the environmental context and the activity of the participant during the recording are well controlled. The task of the participant during the

recording is the critical aspect providing psychological meaning to the physiology. Control of the context implies paying attention to three items: the place, the participant, and the task.

5.1. The Place

The place refers to the physical and social environment of the recording. There are numerous physical and social environmental variables that may affect the physiological recording. The main physical variables are noise, light, temperature, and humidity. The main social variables are the presence of people during the recording (including the experimenter). To control these variables, specific rooms in which such variables could be eliminated or maintained constant should be used. The optimal place is the psychophysiological laboratory, with a participant's room independent of the experimenter's room (Fig. 1).

5.2. The Participant

The participant always introduces a set of variables that affect the physiological recording. Some of these variables are characteristics of the participant such as age, gender, menstrual cycle (in women), whether or not the participant smokes, and whether or not the participant has any physical disease or psychological condition. Other variables affect the participant indirectly such as physical activity done just before the recording, hour of the day, and food, medication, or drinks taken before the recording. All of these variables should be included within the information protocol obtained at the beginning of the assessment session. When possible, they should be controlled either by exclusion of participants with characteristics that saliently affect the recording or by elimination or matching of the conditions affecting the recording.

5.3. The Task

The tasks that should be used depend on the aim of the assessment. Each research area of psychology has developed its own assessment paradigms and tasks. Therefore, assessment must be linked to the theoretical contexts in which the different tasks have been developed. Tasks developed to assess attention are different from tasks developed to assess emotion, memory, language, thinking, motor control, or learning. However, all psychophysiological tasks share a common structure: a period of recording with no stimulation (resting period) used to stabilize the physiological measures and a period of recording with stimulation (the specific task).

6. ASSESSMENT OF BASIC PSYCHOLOGICAL PROCESSES

Interest in psychophysiological assessment has existed since the early years of scientific psychology. Examples of psychophysiological applications for the study of basic processes can be found in the work of James (emotion), Pavlov (learning), and Cannon (motivation). This section presents some examples of current basic applications.

6.1. Electromyography

Electromyography has been used to study arousal, facial expressions, somatic correlates of thought and mental imagery, motor reflexes, motor skill learning, and so forth. An example of current application of EMG to the assessment of emotional and attentional processes is the startle probe paradigm. In humans, the first and most reliable component of startle is eyeblink. Eyeblink is measured using integrated EMG of the orbicularis oculi (i.e., the muscle around the eye). Startle is typically elicited by a brief sudden noise. Emotional modulation of startle is observed when the reflex is elicited while the participant views affective pictures presented for 6 seconds. Compared with neutral pictures, unpleasant pictures increase the reflex and pleasant pictures inhibit it. Attentional modulation of startle is observed when another stimulus precedes the eliciting stimulus by less than 500 milliseconds. Under these conditions, the reflex is inhibited (i.e., prepulse inhibition) independently of the affective content. Normal and pathological emotional and attentional processes can be assessed using this paradigm.

6.2. Heart Rate

Cardiovascular activity has been used in numerous research contexts, including unconditioned reflexes (i.e., orienting and defense), habituation, sensitization, Pavlovian conditioning, instrumental learning, motivation, emotion, memory, and mental workload. Heart rate has been the cardiovascular measure most frequently used in these contexts.

The psychological meaning of heart rate has been a subject of debate in the past. Two major interpretations

can be identified: the cognitive and motivational interpretations. The cognitive interpretation assumes that cardiac changes in response to external or internal stimuli reflect attentional and perceptual mechanisms aimed at facilitating (heart rate deceleration) or inhibiting (heart rate acceleration) sensory processing. On the other hand, the motivational interpretation assumes that cardiac changes reflect metabolic mechanisms aimed at providing energy to adaptive behaviors. If the appropriate behavior is to be passive, the cardiac response will be a heart rate deceleration. If the appropriate behavior is to be active either physically or psychologically, the cardiac response will be a heart rate acceleration.

6.3. Skin Conductance

Skin conductance is one of the most sensitive measures of mental activity. It has been used extensively to study attention, habituation, and conditioning. A skin conductance response to stimuli of moderate or low intensity is interpreted as an attentional response indicative that a novel stimulus has been detected and processed. Habituation of the response (when the same stimulus is repeated) and dishabituation (when the repeated stimulus changes in some parameter) is also indicative that the system is correctly processing the environmental information. In the context of conditioning studies, the skin conductance response is interpreted as indicative that a neutral stimulus has acquired the signal value of the unconditioned stimulus, allowing assessment of the degree of learning or unlearning.

Skin conductance has also proved to be sensitive to unconscious presentation of threat stimuli. Using masked stimuli (visual pictures presented for 30 milliseconds followed by a mask), Öhman and colleagues showed that fear-relevant stimuli (pictures of snakes, spiders, and fear–angry faces), as compared with irrelevant stimuli (pictures of flowers, mushrooms, and neutral–happy faces), can be processed without awareness.

6.4. Central Electromagnetic Measures

Among the central electromagnetic measures, the ERPs are the most frequently used to assess cognitive functions. The N_{100}, P_{300}, N_{400}, and mismatch negativity all are ERPs that are studied extensively in relation to attentional and linguistic processing. Each ERP requires, in addition to the EEG recording, the use of specific tasks.

One of the tasks to obtain the N_{100} and the P_{300} is the oddball task. It consists of a large series of presentations of two stimuli with different degrees of probability, for example, a tone of 1000 hertz presented 160 times (80% probability) and a tone of 1500 hertz presented 40 times (20% probability) in random order. The task of the participant is to pay attention only to the tone with lower probability. Under these conditions, the N_{100} reflects the nonspecific attention paid to both stimuli, whereas the P_{300} reflects the specific attention paid only to the low-probability stimulus.

6.5. Central Metabolic Measures

PET and fMRI are the latest techniques used in psychophysiological assessment. Identification of the brain areas involved in the psychological functions activated by the task requires application of the subtraction method. It consists of obtaining images at different stages of the task. For example, in a semantic processing task in which a participant is asked to think of the meaning of a word, the task is divided into several stages: to hear the word, to hear the word and repeat it, and to hear the word and generate a verb representing a use of the word (e.g., hammer: to hit). By subtracting the images from the different stages, the specific brain areas involved in the act of thinking can be identified.

7. CLINICAL APPLICATIONS

Psychophysiological assessment has been used in many applied areas, including ergonomy, sport, lie detection, and clinical psychology. This section presents some examples of applications within clinical psychology.

7.1. Psychopathological Disorders

Psychopathological disorders are characterized by symptoms of psychological dysfunction with no diagnosable brain lesion or physical disease. Psychophysiological assessment of psychopathology has focused on identification of psychophysiological markers, that is, deviation in one physiological variable that is reliably associated with a specific disorder. For example, research on schizophrenia has suggested several markers, including reduced P_{300}, reduced skin conductance orienting response, reduced prepulse

inhibition of startle, and abnormal eye movements. Similar suggestions have been made in relation to other disorders such as psychopathy, depression, and anxiety, although more research is needed before useful clinical protocols can be developed for these disorders.

7.2. Psychophysiological Disorders

Psychophysiological disorders are characterized by symptoms of physiological dysfunction in specific organs such as the heart, lungs, stomach, muscles, and neural tissues. Such dysfunction is caused or aggravated by psychological factors such as emotions, stress, personality, and lifestyle.

Therefore, psychophysiological assessment is a key procedure for reliable diagnosis of these disorders. Psychophysiological disorders in which an etiological role played by psychological factors has been demonstrated include cardiovascular disorders (e.g., essential hypertension, cardiac arrhythmias, migraine), respiratory disorders (e.g., asthma, hyperventilation), gastrointestinal disorders (e.g., ulcer, irritable colon, colitis, chronic constipation), neuromuscular disorders (e.g., tics, tremblers, tension headache), dermatological disorders (e.g., eczema, acne, psoriasis), sexual disorders (e.g., dysfunction of sexual desire, excitation, and orgasm), and neurological disorders (e.g., sleep disturbances, epilepsy).

7.3. Neuropsychological Disorders

Neuropsychological disorders are characterized by symptoms of psychological dysfunction caused by a diagnosable brain lesion. The psychological dysfunction is usually a cognitive one such as altered perception (agnosias), memory (amnesias), language (aphasias), or motor coordination (apraxias). The neural lesion is usually a stroke, a traumatic brain lesion, a tumor, or a degenerative neural disease.

Although neuropsychological assessment does not normally use psychophysiological methods, recent attempts to combine both methods are promising. For example, ERPs are being used to assess cognitive deficits in Alzheimer's disease, epilepsy, alcoholism, and mental retardation. Similar applications of functional neuroimaging techniques are helping to better understand the specific brain dysfunctions produced by different lesions.

Acknowledgments

The preparation of this article was partially supported by Research Grant BSO2001-3211 from the Spanish Ministry of Science and Technology and by the Research Group HUM-388 of the Junta de Andalucía.

See Also the Following Articles

Clinical Assessment ■ Neuropsychological Assessment in Schools ■ Psychological Assessment, Standards and Guidelines for ■ Psychological Rehabilitation Therapies

Further Reading

Cacioppo, J. T., Tassinary, L. G., & Berntson, G. G. (Eds.). (2000). *Handbook of psychophysiology*. Cambridge, UK: Cambridge University Press.

Cuthbert, B. N., Lang, P. J., Strauss, C., Drobes, D., Patrick, C. J., & Bradley, M. M. (2003). The psychophysiology of anxiety disorder: Fear memory imagery. *Psychophysiology*, *40*, 407–422.

Gazzaniga, M. S., Ivry, R. B., & Mangun, G. R. (Eds.). (2002). *Cognitive neuroscience* (2nd ed.). New York: Norton.

Graham, F. K. (1979). Distinguishing among orienting, defense, and startle reflexes. In H. D. Kimmel, E. H. Van Olst, & J. F. Orlebeke (Eds.), *The orienting reflex in humans* (pp. 137–167). Hillsdale, NJ: Lawrence Erlbaum.

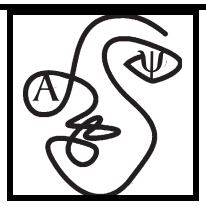
Lang, P. J. (1995). The emotion probe: Studies of motivation and attention. *American Psychologist*, *50*, 372–385.

Obrist, P. A. (1981). *Cardiovascular psychophysiology: A perspective*. New York: Plenum.

Öhman, A., Hamm, A., & Hugdahl, K. (2000). Cognition and the autonomic nervous system: Orienting, anticipation, and conditioning. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of psychophysiology* (pp. 533–575). Cambridge, UK: Cambridge University Press.

Öhman, A., & Soares, J. J. F. (1994). Unconscious anxiety: Phobic responses to masked stimuli. *Journal of Abnormal Psychology*, *103*, 231–240.

Posner, M. I. (1999). *Education and the human brain*. Granada, Spain: University of Granada Press.



Psychotherapy in Older Adults

Bob G. Knight and Steven David

University of Southern California, Los Angeles, California, USA

1. Presenting Problems
2. Assessment, Diagnosis, and Prevalence
3. Interventions
4. Diversity
5. Conclusions
   Further Reading

GLOSSARY

cohort based A focus on generational groups rather than age groups; examples include the World War II generation, the baby-boomer generation, and Generation X.

contextual The idea that changes in therapy often relate to the social–environmental context within which older adults reside rather than their developmental stage; this is especially apparent in settings such as hospitals and nursing homes.

contextual, cohort-based, maturity, specific challenge (CCMSC) model Model that proposes that older adults are more mature than younger adults in certain important ways and yet are concurrently facing some of life's most difficult challenges; the model also asserts that the current social context and past sociocultural circumstances experienced as members of earlier born cohorts must be considered when conceptualizing psychotherapy with older adults.

loss–deficit model A negative conceptual model of aging that portrays late life development as the accumulation of deficits and a series of losses as the normative experience of older adults; the result is characterized by negative mental health outcomes.

maturity The developmental maturation that occurs during normal aging; although this maturation includes changes such as cognitive slowing and use of simpler language, it also includes greater emotional complexity and a greater source of expertise from which to draw.

specific challenges The idea that the high prevalence of chronic medical problems and neurological disorders among older adults will result in the frequent presentation of these problems in assessment and therapy; grief for deceased loved ones and caregiving for other older adults (e.g., spouses, parents) are also specific challenges faced by older clients.

This article integrates discussion of psychotherapy with current research and clinical observations in the field of gerontology. It briefly summarizes past and current research on aging in several areas of practical interest to psychotherapists seeking to apply various psychotherapy treatment techniques in working with older adult clients. The framework for this discussion is the contextual, cohort-based, maturity, specific challenge (CCMSC) model, which grew from the application of gerontology to the practice of psychotherapy with older adults. This model is intended to be nonspecific with regard to system and modality of therapy and, thus, affords a wide utility. The article discusses common presenting problems, issues of assessment and diagnosis, and prevalence rates relevant to psychotherapy with older adults. It reviews the literature concerning effective psychotherapeutic interventions across systems and

©2004 Elsevier Inc.
All rights reserved.

modalities in an effort to report the state of the art. It also addresses issues of diversity and the importance of cultural competence among psychotherapists who treat older adults.

1. PRESENTING PROBLEMS

Older adults seeking therapy often are dealing with problems that would threaten psychological well-being at any age in the life span. These problems include dealing with chronic illness and disability, grieving the loss of loved ones, and caregiving for loved ones with cognitive or physical disability. Although these problems are not unique to later life, they are more likely during this period. Additional problems seen earlier in the life span, such as interpersonal relationship difficulties and personal failures, appear later in the life span as well. Finally, many people who have struggled with depression and anxiety in earlier life continue to struggle with these problems in later life.

Working with distressed older adults often means working with those who are struggling to adjust to the problems associated with being chronically ill or physically disabled. Thus, therapists must acquaint themselves with chronic illnesses and their psychological impact, pain control, adherence to medical treatment, strategies of rehabilitation, and assessment of behavioral signs of medication reactions. The increased presence of chronic illness and disability with each decade of life makes it necessary for therapists working with older adults to develop the ability to discuss physical problems and to discuss when a problem may have physical causes.

Similarly, working with older adults often involves grief work. Although the death of loved ones can be experienced at any point in the life span, this experience is more common in later life. Older adults presenting with depression often have experienced the death of loved ones during recent months or years. In addition, many older adults do not experience the death of loved ones as normal or expected. This is the case even when their loved ones have been ill for some time. These losses are still experienced as surprising and tragic, and they may be experienced more deeply due to the length of their relationships. As a result, psychotherapy with older adults may often require grief work.

Finally, older adults often find themselves in the role of caregivers for physically or cognitively impaired spouses or parents. The course of these physical and cognitive disorders is frequently progressive and debilitating. As such, older adult caregivers are forced to adapt to increasingly difficult caregiving tasks and, consequently, may begin to appraise caregiving as a burden. Thus, therapists working with older adult caregivers must be prepared to treat the stress resulting from this burden.

2. ASSESSMENT, DIAGNOSIS, AND PREVALENCE

Prevalence studies of depression in older adults who are chronically ill or physically disabled have reported rates up to 59%. Further examination of the relationship between depression and physical illness finds that the disability often accompanying physical illness, rather than the illness itself, is a risk factor for depression in older adults. Conversely, depression can be seen as a serious physical and psychological health problem because it has been found to be associated with increases in disability, poorer rehabilitation from physical health problems, and greater risk of mortality. Physicians frequently do not detect symptoms of depression in their older adult clients, and symptoms of depression in older adults often go untreated. This problem is exacerbated by the fact that older adults often present symptoms of depression to their primary care physicians rather than attempting to seek out mental health professionals.

Depression can also be a problem for older adults suffering from Alzheimer's disease or other forms of dementia. Thus, the issue of screening clients with depression for dementia is an important one when working with older adults. Conversely, because depression and dementia may present similar symptoms, clients suspected of having dementia should be screened for depression. In addition, clinicians should be aware of possible cognitive impairment and should include a brief screening instrument, such as the Folstein Mini-Mental State Exam, as a matter of routine when assessing older adult clients. Signs of cognitive impairment on initial screening indicate the need for more extensive cognitive assessment and may necessitate adjustment of treatment goals and intervention style.

Although depression is prevalent in older adults who experience chronic illness with disability, overall research contradicts the stereotype of the older adult as frequently depressed. In reality, depression is less

prevalent in older adults than in younger adults. However, it is important to note that although prevalence of major depressive disorder appears to be low among older adults, studies of self-reported depressive symptoms show increases among older adults. As such, care should be taken to assess both clinical and subclinical depression in older adults. Similarly, multiple symptoms of anxiety may exist among older adults even in cases where a constellation of symptoms meeting diagnostic criteria for anxiety disorders is not present. Regarding older adults who do indeed meet diagnostic criteria, the combined prevalence of all anxiety disorders in adults over 65 years of age is approximately 5.5%.

Although prevalence rates of alcoholism for older adults are lower than those for younger adults, there remain many serious concerns involving alcohol and older adults. In addition to the myriad of psychosocial ramifications of alcohol addiction, other problems include increased risk of falling, potential liver damage, adverse medication interactions, and decreased cognitive functioning. Although some older adult alcoholics have been addicted since younger adulthood, others develop problems with alcohol in older adulthood in response to inadequate social support, depression, and loneliness. This occurs at higher rates for older adults than for younger adults.

Finally, insomnia should be noted as a common problem among older adults. Studies have found that 10 to 25% of adults over 65 years of age report experiencing chronic sleep difficulties. Sleep disturbances may influence daily functioning and, thus, should be included as part of standard assessment practice with older adults.

3. INTERVENTIONS

3.1. Cognitive–Behavioral Therapy

Empirical evidence regarding the effectiveness of psychological interventions with older adults has been collected over the past two to three decades. In a 1994 meta-analysis by Scogin and McElreath of mainly cognitive–behavioral psychological interventions in the treatment of depression in older adults, an aggregate effect size was shown to be roughly equal to that found in another 1994 meta-analysis by Schneider for the efficacy of antidepressant medications.

The factors affecting outcomes with older adults were explored in a 2001 meta-analysis by Pinquart and Sörensen. These include psychotherapy as superior to psychosocial interventions, more than nine sessions of therapy, and specialized training of the therapist. Older adults with clinical depression improved more than those considered to be at risk for depression. It was found that individually based interventions for both depression and caregiver distress were generally more efficacious than group-based interventions. A 2003 study by Wetherell and colleagues of treatment of generalized anxiety disorder found that brief cognitive–behavioral therapy (CBT) had large beneficial effects, and a discussion group treatment had medium beneficial effects, among older adults with generalized anxiety disorder. These results support the effectiveness with older adults of professional therapy by well-trained providers.

In an effort to establish a set of standards for evaluation of therapy effectiveness, the American Psychological Association's Division 12 has developed guidelines to identify empirically supported treatments for specific psychological problems. In 1998, Gatz and colleagues used the then current guidelines in their evaluation of the literature on psychological treatments with older adults. They reported that efficacious treatments for older adults included CBT for sleep disorders and cognitive and behavioral treatments for clinical depression. Gatz and colleagues also found strong support for the effectiveness of behavioral interventions and environmental modifications for persons with dementia.

With regard to subclinical problems in later life, Gatz and colleagues also found that CBT had above-average effects on self-ratings of subjective well-being. In support of these findings, Pinquart and Sörensen's meta-analysis of psychotherapeutic interventions with older adults found that CBT had above-average effects on depression and other self-ratings of subjective well-being. In addition, Gatz and colleagues reported that for nonsyndromal problems of aging, cognitive training is often efficacious in slowing cognitive decline. Therefore, there is considerable evidence for the effectiveness of CBT with older adults.

3.2. Other Individual Therapies

The use of a short-term manualized form of psychodynamic psychotherapy (PDP) has also been studied among older adults. In a 1982 study comparing CBT with PDP among depressed older adults, Gallagher and Thompson found that PDP and CBT were equally effective at posttest. However, this effect did not hold up

for PDP at follow-up. Findings from Schneider's study support the use of a PDP variation, interpersonal counseling (IPC), with older adults. Gatz and colleagues' review of empirically validated treatments examined several studies that included brief psychodynamic therapy as a treatment condition for depression among older adults and found it to be equally as efficacious as CBT. As such, brief psychodynamic therapy can be considered an often efficacious treatment for depression in late life. In addition, Gatz and colleagues concluded that the technique of structured life review or reminiscence is often efficacious for treating depressive symptoms among older adults. However, it is noted that several studies suggest that when life review is compared with other methods of treatment, the alternative methods are more efficacious.

3.3. Family and Caregiving Interventions

Some of the most salient issues in psychotherapy with older adults appear within the family context. Older adults often depend on family members for both emotional and instrumental support. In the event that these family relationships become strained, disruption of support can ensue, resulting in distress for older adults. Exploration of older adults' family histories and interpersonal relationships can often inform the therapeutic process in such cases. In 2000, Qualls argued that although few empirical studies examining the efficacy of family therapy for older adults have been conducted, a solid rationale for its use with this population exists. Prominent in this rationale is the assertion that the lives of many older adults are essentially embedded within the functional context of family. Similarly, the contextual, cohort-based, maturity, specific challenge (CCMSC) model can be applied to the adaptation of family therapy for older adults. This adaptation includes the examination of specific challenges facing older relatives and the family system, the integration of cross-cohort views when working across a multiple-generation family structure, and the influence of age-related social context on older adult family members.

Sörensen and colleagues' meta-analysis of interventions with caregivers of older adult family members with physical or dementing illnesses found that these interventions were generally effective in reducing burden and depression, increasing subjective well-being, and increasing caregiver mastery. It is important to note that although these results were significant, the effect sizes were, on average, small to moderate. These appear to be lower than the previously mentioned effect sizes for the treatment of depression among older adults. In that study, psychoeducational and psychotherapeutic interventions consistently produced short-term effects on outcomes. The majority of these effects remained intact after an average of 7 months postintervention. It was also found that intervention effects were smaller for dementia caregivers than for caregivers of family members with other illnesses.

4. DIVERSITY

4.1. Ethnicity

The CCMSC model asserts that the majority of diversity issues in psychotherapy with older adults can be considered as cohort and contextual differences among clients. These differences are apparent when ethnicity, gender, and sexual identity are considered. Cohort differences are particularly salient in relation to ethnicity. Ethnic minorities belonging to earlier born cohorts have had to contend with especially difficult experiences of racism and economic hardship. In addition, social context may vary by ethnicity. In some ethnic groups, the use of large intergenerational family structures and an increased emphasis on familism is prominent. It is important for clinicians to become sensitive to potential ethnicity-related contextual or cohort differences. Some of these differences may include verbal and nonverbal behavior and communication styles that differ according to the level of acculturation. Thus, rather than making assumptions about clients, it is incumbent on clinicians to respectfully encourage clients to openly define for themselves the salience of their ethnicity to their issues in therapy.

4.2. Gender

In addition to contextual and cohort differences related to ethnicity, gender-related differences should be considered. The majority of older adults are women, and women tend to be overrepresented among therapy clients. Earlier born cohorts of women have seen the metamorphosis of women's roles in society over time, yet many may find that their personal experiences differ from those of later born cohorts. Many of these women may approach therapy with perceptions that are unique to their cohorts. In addition, because many

women find themselves surviving their husbands in later life, they may suddenly be faced with role transitions that are presenting themselves for the first time. Therefore, clinicians must be sensitive to the unique challenges faced by women in later life.

4.3. Sexual Identity

Finally, differences related to sexual minority status should be considered. Cohort differences among lesbian, gay, bisexual, and transgender (LGBT) older adults reveal differences between earlier born cohorts coming of age prior to the 1969 Stonewall riots (and the gay rights movement that followed) and those cohorts whose members matured later. However, some of these LGBT older adults may have been involved in earlier smaller movements that supported rights for sexual minorities. Thus, although many of these earlier born cohorts may have experienced a large portion of their lives under the weight of secrecy and shame, others may have experienced a more publicly expressed identity. Nevertheless, a range of LGBT older adults have endured the tremendous stress of both personally and socially experienced stigmatization. As with ethnicity, it is important for clinicians to refrain from developing assumptions regarding LGBT older adult clients. Instead, clinicians should allow these clients to define for themselves the salience of their sexual minority status to issues in therapy.

5. CONCLUSIONS

The CCMSC model proposes that older adults are more mature than younger adults in certain important ways and yet are concurrently facing some of life's most difficult challenges, including adjusting to chronic illness and disability in addition to frequently grieving the loss of others. The model also asserts that the current social context and past sociocultural circumstances experienced as members of earlier born cohorts must be considered when conceptualizing psychotherapy with older adults. Older adults seeking therapy often are dealing with problems that include coping with chronic illness and disability, grieving the loss of loved ones, and caregiving for loved ones with cognitive or physical disability. Although these problems are not unique to later life, they are more frequent during this period. Prevalence studies indicate the presence of depression and anxiety disorders among older adults, albeit at lower rates than those among younger adults. Nevertheless, it is important to note that although rates of clinical depression decline in later life, rates of self-reported symptoms of depression increase. Thus, clinicians should attend to the possible presence of subclinical depression among older adults. In addition, issues of alcoholism and sleep disorders appear during older adulthood.

Research on empirically validated treatments has shown that efficacious interventions among older adults include (a) CBT, IPT, brief psychodynamic therapy, and life review for depression; (b) CBT for increases in subjective well-being; (c) CBT for sleep disorders; (d) cognitive treatment to slow nonsyndromal cognitive decline; and (e) psychoeducational and psychotherapeutic interventions for caregiver distress.

Finally, in keeping with the CCMSC model, contextual and cohort-based issues should be considered regarding the diverse experiences of ethnicity, gender, and sexual identity among older adults. Clinicians must be sensitive to this diversity by setting aside assumptions and encouraging the expression of diversity throughout clients' therapeutic process.

Acknowledgment

The second author was supported in part by a National Institute on Aging multidisciplinary training grant.

See Also the Following Articles

Age-Related Issues among Minority Populations ■ Aging and Competency ■ Anxiety Disorders in Late Life ■ Depression in Late Life ■ Personality and Emotion in Late Life

Further Reading

Gallagher, D. E., & Thompson, L. W. (1982). Treatment of major depressive disorder in older adult outpatients with brief psychotherapies. *Psychotherapy: Theory, Research, and Practice*, *19*, 482–490.

Gatz, M., Fiske, A., Fox, L. S., Kaskie, B., Kasl-Godley, J. E., McCallum, T. J., & Wetherell, J. L. (1998). Empirically validated psychological treatments for older adults. *Journal of Mental Health and Aging*, *4*, 9–46.

Knight, B. G. (2004). *Psychotherapy with older adults* (3rd ed.). Thousand Oaks, CA: Sage.

Pinquart, M., & Sörensen, S. (2001). How effective are psychotherapeutic and other psychosocial interventions with older adults? A meta-analysis. *Journal of Mental Health and Aging*, *7*, 207–243.

Qualls, S. H. (2000). Therapy with aging families: Rationale, opportunities, and challenges. *Aging & Mental Health*, *4*, 191–199.

Schneider, L. S. (1994). Meta-analysis from a clinician's perspective. In L. S. Schneider, & C. F. Reynolds, III (Eds.), *Diagnosis and treatment of depression in late life: Results of the NIH Consensus Development Conference* (pp. 361–373). Washington, DC: American Psychiatric Association.

Scogin, F., & McElreath, L. (1994). Efficacy of psychosocial treatments for geriatric depression: A quantitative review. *Journal of Consulting and Clinical Psychology*, *62*, 69–73.

Sörensen, S., Pinquart, M., & Duberstein, P. (2002). How effective are interventions with caregivers? An updated meta-analysis. *The Gerontologist*, *42*, 356–372.

Wetherell, J. L., Gatz, M., & Craske, M. G. (2003). Treatment of generalized anxiety disorder in older adults. *Journal of Consulting and Clinical Psychology*, *71*, 31–40.

Racial and Ethnic Minorities, Counseling of

Edward A. Delgado-Romero
Indiana University, Bloomington, Indiana, USA

Debra Mollen
Texas Women's University, Texas, USA

Charles R. Ridley
Indiana University, Bloomington, Indiana, USA

1. Introduction
2. Motivation for Counseling Racial and Ethnic Minorities
3. Acculturation and Counseling
4. Unintentional Racism/Aversive Racism
5. Racial Identity
6. Multicultural Counseling Competence
Further Reading

GLOSSARY

acculturation A process of attitudinal and behavior change experienced by individuals who have come in contact with a different and dominant culture.
competence The ability to predetermine and then to attain the desired outcomes of one's profession or vocation by coordinating and integrating relevant skills and microskills.
counseling The purposeful development and structuring of a professional relationship of which the intention is to facilitate meaningful changes in the lives of its recipients. The focus is typically on normal and developmental aspects of functioning as opposed to psychopathology.
culture The ideals, beliefs, skills, customs, values, and social behavior of a particular group of people.
ethnicity A group of humans identified as distinctive by cultural practices and/or beliefs.
multicultural Relative to two or more groups or cultures.
race A group of humans identified as distinctive by physical characteristics such as skin color.
racial identity The view one has about the role of race in dictating social group membership.
racism Fear, hatred, and oppression solely based on the perceived racial or ethnic characteristics; may be intentional or unintentional.

Counseling racial and ethnic minorities entails the provision of counseling services (including individual and group therapy, consultation, and research) to members of racial and ethnic minorities in the United States. In response to societal and demographic changes, as well as a growing ethical awareness, the need to address completely the unique needs of racial and ethnic minorities has been a powerful influence in counseling psychology. Of particular importance are the relationship of acculturation to counseling, the influence of unintentional or aversive racism in counseling, racial identity in regard to counseling, and especially the overarching issue of what constitutes multicultural counseling competence.

1. INTRODUCTION

Historically, the field of applied psychology has been negligent in its delivery of services to racial and ethnic

©2004 Elsevier Inc.
All rights reserved.

minorities. The assumptions that have guided practice throughout the history of psychology have relied on White (English-speaking, European-descended, heterosexual) males as the standard for normalcy. Persons from other backgrounds (i.e., race, ethnicity, socioeconomic status, and sexual orientation) were conceptualized, regarded, and treated as deviants; systematically denied access to quality services; and often discounted and ignored by psychology professionals. As the United States entered the Civil Rights era in the 1950s, the injustice of ignoring race and culture in counseling psychology became apparent. In 1962, Wrenn coined the phrase "cultural encapsulation" to describe counselors who defined a client's reality according to their own cultural assumptions and stereotypes, minimized cultural differences in favor of universal explanations, disregarded their own cultural biases, conceptualized other cultures in terms of pathology, and used a technique-oriented model of the counseling process.

Gradually, partially in response to the Civil Rights movement, applied psychology grew more reflective as a specialty. Practitioners, researchers, and scholars began to recognize the extensive harm caused by the profession's racial and ethnic biases and prejudices. Professionals began to observe and identify biases in their own approach to minority clients. It was clear that applied psychology was far from being a value-neutral profession.

In what has commonly been referred to as psychology's "fourth force," multiculturalism has assumed a position of significant impact in the past two decades, prompting psychologists to revisit and repair their damaging beliefs and practices as well as to make a concerted effort to ensure that counseling services are accessible and amenable to racial and ethnic minority clients. Concurrent with the new awareness has come a growing commitment to training members of racial and ethnic minority populations as counselors.

Progress in the field has been hampered by unclear definitions of the constructs of race and ethnicity and vague descriptions of who constitutes a minority. Race and ethnicity have been criticized as being unscientific and socially constructed categories that are inadequately defined and operationalized. Often, issues of race and ethnicity are confounded by other factors, such as socioeconomic status. In addition, the term multicultural has been criticized as too general and serving to obscure unresolved racial issues.

2. MOTIVATION FOR COUNSELING RACIAL AND ETHNIC MINORITIES

In addition to the historical and societal forces identified previously, two main motivations for the competent counseling of racial and ethnic minorities have been identified: demographic change and ethical concerns. In terms of demographics, the United States has seen a steady increase in some minority populations. The U.S. Census has documented the changing demographic composition of the United States. Census predictions suggest that by 2050 Whites will no longer represent a majority of the U.S. population; in some states, such as California, this change is already a demographic reality. The increase of racial and ethnic minorities, however, does not automatically yield economic success, access to education, or political empowerment. On the contrary, although in some areas people of color are no longer a statistical minority, true economic, educational, and political equality for them remains an ideal to which many aspire rather than a reality that is practiced.

Both psychology's and counseling's ethical guidelines have reflected the evolution toward a more culturally sensitive applied profession, reflecting the theme of social justice. The most recent version of each set of guidelines takes culture—defined broadly to include age, gender, race, ethnicity, national origin, religion, sexual orientation, disability, language, and socioeconomic status—into account. The guidelines impart the expectations that counselors and psychologists be adequately prepared to treat clients from diverse backgrounds, that they do not practice unfair discrimination, and that they consider culture in the practice of applied psychology. However, professional ethics codes and standards of conduct—believed by many to be an improvement over the earliest versions, in which consideration of culture was absent—still meet with criticism by many scholars. Critiques tend to focus on the underpinnings of the codes, which are individualistic, based on principle ethics, and therefore problematic when viewed through a collectivistic lens; on the vagueness of the standards and practices, which fail to give practitioners sufficient direction; and on the absolutist assumption, which fails to give culture satisfactory emphasis. Considerable work must be done before the profession can ensure that racial and ethnic minority clients receive ethically appropriate treatment from culturally sensitive, competent practitioners.

3. ACCULTURATION AND COUNSELING

Acculturation can be defined as a process of attitudinal and behavioral change experienced by individuals who live in multicultural societies or who have come in contact with a different culture due to colonization, invasion, political change, globalization, and the increased mobility of society due to technological advances. Acculturation refers to the balance between changing attitudes and behaviors as a result of contact with a dominant group and retention of existing cultural values, beliefs, and traditions. In the United States, institutions such as the educational system and the media are a part of the acculturation process. Another component of the acculturation process may include the acceptance of the need for counseling. The degree of acculturation to a dominant culture may also play a role in preference for the ethnicity of one's counselor or the client's willingness to discuss racial, ethnic, or cultural issues. For example, a client who has low levels of acculturation may not view counseling as an option and instead prefer to rely on family or other culturally sanctioned sources of support. On the other hand, a highly acculturated individual may be very receptive to counseling and may even prefer to work with a White counselor.

4. UNINTENTIONAL RACISM/ AVERSIVE RACISM

In a cultural zeitgeist of increased awareness about the damaging effects of racial and ethnic prejudice and discrimination, overt acts of racism have abated to some extent. It would be erroneous to conclude, however, that decreased incidents of blatant racism indicate an absence of racial and ethnic prejudice. Instead, as public and professional awareness about racism increases, more covert examples of racism are evident. These have been termed aversive and unintentional racism.

Aversive racism occurs when individuals who identify themselves and others as egalitarian and unprejudiced subtly discriminate against racial and ethnic minorities in their beliefs and feelings. This may occur when such prejudice can be rationalized or when information the individual has received is ambiguous. In 1995, Ridley identified eight methods of unintentional but pervasive racist defenses that occur in counseling. Color blindness is the tendency for counselors to disregard the race of their clients, assuming that all clients are similar. They may minimize or deny the harm inherent in racism, prejudice, and discrimination. Color consciousness, on the other hand, occurs when counselors place too much emphasis on race, ignoring pathology or attributing all presenting problems to culture. Cultural transference can occur when racial and ethnic minority clients project their feelings about the majority culture onto their counselor, whereas cultural countertransference indicates the counselor's projection of emotional reactions toward a minority group onto a particular minority client. Cultural ambivalence can occur when counselors use the therapy arena to act out their own feelings about race and ethnicity, either by dominating and controlling or by seeking approval and acceptance from the client. Pseudo-transference refers to the tendency for counselors to discount their clients' perceptions of them as racist and prejudiced; the clients' reactions to their counselors may be casually dismissed as transference when actually rooted in reality. Overidentification occurs when minority counselors relate significantly with their minority clients and may unwittingly focus on racial issues to the neglect of other important presenting problems. Finally, identification with the oppressor refers to the tendency for some racial and ethnic minority counselors to deny their experiences as minorities and overidentify instead with the majority culture. Ridley offered counselors a variety of ways to combat unintentional racism.

5. RACIAL IDENTITY

The ways individuals perceive themselves as racial beings form the processes by which people acquire their racial identity, and thereby influence both the counseling process and the selection of counselors. Researchers have proposed several different models that outline the distinctive processes that formulate racial identity.

Phinney based her 1993 model of ethnic identity development primarily on her work with adolescents. She proposed a three-stage model to capture the process by which adolescents formulate their ethnic identity. The first stage, unexamined ethnic identity, represents a relatively static condition in which individuals are not in the process of exploring themselves as

ethnic individuals. In the second stage, ethnic identity search/moratorium, individuals are involved in actively exploring themselves ethnically and trying to make meaning of their ethnicity. In the third stage, achieved ethnic identity, individuals have a clear understanding of themselves in terms of their ethnic identity.

Updating Cross's earlier work on Black identity, Cross and Vandiver in 2001 proposed a four-stage model to elucidate racial identity formulation, referred to as the expanded nigrescence model. In stage 1, preencounter, individuals have not yet begun to explore themselves as racial beings and may even denigrate their own culture. Eventually, individuals encounter an experience that prompts them to begin to explore themselves as racial people; this experience is referred to as an encounter experience. This can include an incident involving racism that is experienced personally or having a powerful response to a cultural event, historical incident, or other racial experience. In stage 2, immersion–emersion, individuals experience the most pronounced period of transition. As they continue to concretize their racial identity, they may embrace all things affiliated with their racial or ethnic group and they may eschew those things affiliated with the dominant culture. In stage 3, internalization, individuals have worked through a significant portion of the transition period and may embrace their racial identities as one facet of themselves while simultaneously considering multiple salient roles. In stage 4, internalization–commitment, individuals sustain their interest in their racial identity and may choose to devote themselves to promoting the needs and interests of their particular ethnic group. Cross and Vandiver note that there are cluster identities within each stage, and these cluster identities acknowledge that within stages individuals will experience identity development differently.

Therefore, according to the stage of their racial identity development, racial minority clients may have different views about the purpose of counseling, the preferred race of their therapist, and what they want to work on in counseling. For example, a client in the preencounter stage may deny the salience of race and endorse "color-blind" views of the United States. A client in the immersion–emersion stage may need to explore actively the meaning of race and struggle to make sense of an encounter experience. However, it is not only the client who undergoes racial identity development; the counselor also has a racial identity and undergoes development. The interaction of the two identities can be a powerful dynamic in therapy. Cross and Vandiver's model may also be generalized beyond Black racial identity and serve as an exemplar of the racial or ethnic identity process.

Janet Helms has utilized her models both on White racial identity and on people of color racial identity to create an interpersonal process model that may be particularly useful for counselors seeking a greater understanding of the dynamics that occur in cross-racial dyads. Helms posited that it is the expressed racial identity of the individuals involved in a particular dialogue that is most critical. In parallel interactions, individuals share similar attitudes and assumptions and generally try to maintain accord. In regressive interactions, the person afforded the most power socially functions from a less evolved status than the person with less power. In progressive interactions, the individual with more social power responds from a relatively more evolved status than the person with less social power. Finally, in crossed interactions, participants interact from places that are diametrically opposite in terms of responding to racial matters.

6. MULTICULTURAL COUNSELING COMPETENCE

During the past two decades, significant attention has been focused on the notion of multicultural counseling competence. The profession of applied psychology has particularly endorsed the need for practitioners who are multiculturally competent.

One such approach originally proposed by Sue and colleagues in 1992 consisted of creating guidelines for multiculturally competent practitioners. This entailed specific directives based on a tripartite model of knowledge, skills, and attitudes. The focus is on the acquisition and endorsement of those facets of competence endorsed by the authors and supported by the profession at large. In this view, multicultural counseling competence consists of professionals who are actively open to examining and adjusting their own assumptions and worldviews, who work toward nonjudgmentally understanding their clients' viewpoints, and who practice appropriate and sensitive strategies with clients from diverse racial and ethnic backgrounds. Many instruments have been created in response to the tripartite considerations of knowledge, skills, and attitudes.

Another approach by Ridley and Mollen recognizes the importance of the aforementioned competencies model and endorses a model that embraces an

integrative deep structure as well—a structure that involves such facets of practice as purposeful application, motivation, selection, sequencing, and timing. This approach posits that multicultural counseling competence consists not only of the building blocks of therapy called microskills and competencies but also of the purposeful, committed intention of the counselors to create an environment suitable for meaningful and relevant client change in consideration of racial, ethnic, and cultural variables.

See Also the Following Articles

Acculturation ■ Conformity across Cultures ■ Cross-Cultural Psychology, Overview ■ Health Psychology, Cross-Cultural

Further Reading

American Psychological Association. (2003). Guidelines on multicultural education, training, research, practice and organizational change for psychologists. *American Psychologist*, *58*, 377–402.

Cross, W. E., & Vandiver, B. J. (2001). Nigrescence theory and measurement: Introducing the Cross Racial Identity Scale (CRIS). In J. G. Ponterotto, J. M. Casas, L. A. Suzuki, & C. M. Alexander (Eds.), *Handbook of multicultural counseling* (2nd ed., pp. 371–393). Thousand Oaks, CA: Sage.

Helms, J. E. (1984). Toward a theoretical explanation of the effects of race on counseling: A black and white model. *The Counseling Psychologist*, *12*, 153–165.

Helms, J. E. (1995). An update of Helm's white and people of color racial identity models. In J. G. Ponterotto, J. M. Casas, L. A. Suzuki, & C. M. Alexander (Eds.), *Handbook of multicultural counseling* (pp. 181–191). Thousand Oaks, CA: Sage.

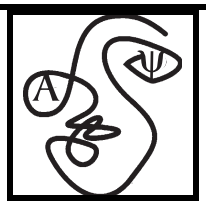
Phinney, J. (1993). A three-stage model of ethnic identity development. In M. Bernal, & G. Knight (Eds.), *Ethnic identity: Formation and transmission among Hispanics and other minorities* (pp. 61–79). Albany: State University of New York Press.

Ponterotto, J. G., Casas, J. M., Suzuki, L. A., & Alexander, C. M. (Eds.). (2001). *Handbook of multicultural counseling* (2nd ed.). Thousand Oaks, CA: Sage.

Ridley, C. R. (1995). *Overcoming unintentional racism in counseling and therapy: A practitioner's guide to intentional intervention* (Multicultural aspects of counseling series, Vol. 5). Thousand Oaks, CA: Sage.

Sue, D. W., Arredondo, P., & McDavis, R. J. (1992). Multicultural counseling competencies and standards: A call to the profession. *Journal of Counseling and Development*, *70*, 477–483.

Sue, D. W., & Sue, D. (1999). *Counseling the culturally different* (3rd ed.). New York: Wiley.



Rape Prevention

Barbara Krahé
University of Potsdam, Potsdam, Germany

1. Rape: Definition and Prevalence
2. Prevention Measures Targeting Individuals
3. Prevention Measures at the Societal Level
4. Conclusion
Further Reading

GLOSSARY

rape Oral, anal, or vaginal penetration through threat or use of force or intentional incapacitation.
rape myth acceptance Tendency to accept as true statements about rape and rape victims that imply victim blame and are not supported by data (e.g., "Every woman has the secret desire to be raped").
rape prevention beliefs Perceptions of personal risk and of the effectiveness of different preventive strategies.
rape shield laws Legal restrictions limiting the admission of evidence about complainant sexual history.
rape trauma syndrome A distinct pattern of psychological and physical symptoms characteristic of victims of sexual assault.
real rape stereotype Idea that the genuine rape is a surprise attack in a dark alleyway by an assailant who is a stranger to the victim and uses force during the assault.

Rape is a pervasive problem in societies around the world and causes severe suffering and trauma to victims. The vast majority of rapes are committed by men against women, but male-on-male rape and female sexual aggression against men have also been recognized during recent years, as reflected in changes in legal definitions as well as in research attention. Therefore, interventions designed to reduce the rate of rape are of paramount significance. Most interventions are directed at individuals, targeting either men as potential perpetrators or women as potential victims. Systematic evaluations of these programs show that their success is limited. In addition, measures are implemented at the societal level to enforce more rigorous legal prosecution of sexual offenses and to challenge traditional gender stereotypes that are conducive to rape.

1. RAPE: DEFINITION AND PREVALENCE

In a worldwide survey of sexual violence published in 2002, the World Health Organization defined rape as physically forced or otherwise coerced penetration—even if slight—of the vulva or anus, using a penis, another body part, or an object. Although this definition provides a common basis for documenting the prevalence of rape across a range of countries, legal definitions of the offense show some variations in the key elements in the definition of rape. In the United States, for example, the Federal Bureau of Investigation (FBI) adopts a definition of rape that is restricted to forcible vaginal intercourse with a woman without her consent by force or threat of force. In contrast, most federal states have adopted legislation in which rape is defined more broadly in terms of oral, anal, or vaginal

©2004 Elsevier Inc.
All rights reserved.

penetration through threat or use of force or intentional incapacitation. There is no longer a restriction of legal definitions to extramarital acts of sexual aggression, thereby instituting "marital rape" as a criminal offense. Similar legal definitions can be found in other Western countries. In Britain, rape is legally defined as sexual intercourse (i.e., vaginal or anal penetration with the penis) without the other person's consent and where the defendant knows that the other person does not consent or is reckless as to whether that person is consenting or not. In Germany, the current legal definition includes the use or threat of violence or the exploitation of the victim's incapacitated state to force another person into sexual activities. Forced sexual activities qualify as rape if the act involves penetration of the victim's body, and this is no longer restricted to penile–vaginal intercourse.

Official crime statistics record only those incidents of rape that are reported to the police, and there is a general consensus that they severely underestimate the true scale of rape. Nonetheless, these sources show that rape is a widespread problem. For 2001, the Uniform Crime Reports reveal that 62.2 per 100,000 women in the United States were victims of forcible rape and that a rape was committed every 5.8 minutes. In terms of lifetime prevalence, it is estimated that 15% of women in the United States become the victims of rape at some point in their lives. Figures in other Western countries are lower but still substantial, for example, a rate of 13.6 per 100,000 female citizens in Germany in 2000 (including both use/threat of force and exploitation of incapacitation).

Evidence based on community surveys is able to detect at least a proportion of cases not reported to the police and, therefore, typically yields higher prevalence rates. In 1987, Koss and colleagues published a groundbreaking survey of college students from 32 higher education institutions in the United States that found that 15.4% of all women reported they had been raped since 14 years of age. In 1999, Spitzberg integrated evidence from 120 studies with a total of more than 100,000 women and arrived at a similar figure of 12.9%. Evidence compiled in the World Health Organization report of 2002 from a range of samples in Africa, Latin American, Asia, and Eastern Europe reflect prevalence rates between 0.3% (The Philippines) and 8.0% (Brazil), but these figures include only incidents from 16 years of age onward.

The majority of rapes are committed by perpetrators known to the victims either as acquaintances or as dating partners. Despite the persistence of the "real rape stereotype" picturing rape as a surprise attack in a dark alleyway by an assailant who uses force during the assault, sexual assaults by strangers are the exception rather than the rule. This fact has important implications for rape prevention efforts. Such efforts can be broadly classified into two groups: (a) prevention efforts targeting individuals as potential perpetrators or victims and (b) measures at the societal level designed to discourage sexual violence.

2. PREVENTION MEASURES TARGETING INDIVIDUALS

The majority of rape prevention efforts are directed at individuals, aiming to make men less likely to engage in sexual violence and to prepare women to anticipate and avoid sexual assault. The specific aims and contents of such measures vary depending on the target audience. Three main types of programs have been implemented and evaluated in systematic research: (a) programs directed at men only, (b) programs directed at women only, and (c) programs directed at mixed-sex audiences.

2.1. Programs Directed at All-Male Audiences

The purpose of these programs is to address men's cognitive and affective predispositions toward sexual aggression. Several studies have demonstrated that greater knowledge about rape, particularly about the rape trauma syndrome, is associated with men's lower likelihood of committing sexually aggressive acts. Measures in this category are directed at changing men's rape-supportive attitudes, such as rape myth acceptance, and promoting information about the adverse effects of sexual assault on the victims. They typically consist of one or two sessions lasting up to 2 hours each. In terms of success, meta-analytic evidence shows that human sexuality courses, workshops, video interventions, and other formats have short-term success in reducing rape myth acceptance. However, to prevent men from acting out rape myths in sexually aggressive behavior, such intervention effects need to persist over a longer period. In addition to preventing first-time sexual assault through supplementing rape-supportive attitudes by greater victim empathy, programs need to focus on sexually aggressive men to prevent them from reoffending. This is particularly

important given the high proportion of multiple offenders among sexually aggressive men. One approach is to train men to discriminate between their arousal by consensual sexual stimuli and that by nonconsensual sexual stimuli and to teach them to activate inhibitory self-persuasion with regard to nonconsensual sexual cues. However, programs directed specifically at men with histories of sexual aggression show that this group responds less well to rape prevention efforts, even in the short term, than do unselected male audiences.

2.2. Programs Directed at All-Female Audiences

This category includes measures directed at women as the most likely victims of sexual assault. Measures are designed to reduce the risk of sexual victimization by alerting women to potentially dangerous situations and behaviors and by enabling them to resist sexual assaults more successfully. Despite the fact that the ultimate responsibility for avoiding rape lies with the perpetrators, there is now a broad consensus that women should be empowered to take a more active self-protective role. This includes both provision of information about the risk factors of sexual victimization and specific advice and training on how to fight off an attack.

High levels of sexual activity, alcohol consumption in the context of sexual encounters, and the ambiguous communication of sexual intentions have been linked to an increased risk of sexual victimization across a range of studies. Women must be informed about the risks involved in these behaviors to give them the opportunity to adjust themselves to the threat of rape. Programs teaching self-defense skills and strengthening women's assertiveness are aimed at reducing women's feelings of vulnerability and increasing their self-confidence when faced with a potentially dangerous situation. However, systematic evaluations of rape prevention programs suggest that their success is limited. In particular, there is evidence showing that rape prevention programs may lead to a small reduction in the incidence of first-time assaults on women but are ineffective in reducing the risk of revictimization in women who had experienced prior sexual assault. In addition, some programs were found to increase participants' knowledge about rape but failed to affect the incidence of sexual victimization compared with a control group.

Whether or not women are willing to take behavioral precautions against the risk of sexual assault depends, to a large extent, on their "rape prevention beliefs," that is, their perceptions of personal risk and of the effectiveness of various preventive strategies. The rape prevention beliefs of most women are restricted to the stranger rape scenario and fail to acknowledge the risk of acquaintance rape. Moreover, rape prevention beliefs are primarily concerned with responding to a sexual attack once is has occurred so as to minimize harm rather than with avoiding an assault in the first place. Women generally perceive a low personal risk of acquaintance rape and, therefore, are less likely to see a need to adopt protective behaviors against this form of threat.

An important component of rape prevention efforts directed at women is to promote women's reflection about potential protective responses in an assault situation. There is conclusive evidence suggesting that fighting back is more likely to succeed in avoiding rape than to lead to greater injury. Therefore, most researchers advise that active self-defense is a prudent response to a sexual attack. However, it is important to bear in mind that the positive evaluation of active resistance is based on probability estimates for samples of assault victims and does not guarantee that resistance is the best strategy in any single case. Therefore, authors are careful to stress that the decision for or against physically fighting back in an attack must be based, in the final instance, on an assessment of the particular circumstances of the attack.

However, it has been noted that empowering women to fight back when they are sexually assaulted involves more than providing self-defense skills. It requires a fundamental challenge of traditional gender stereotypes that portray active physical resistance as unfeminine and, therefore, as inappropriate for women. This applies in particular to rape resistance in assaults by acquaintances or dates where women may be reluctant to acknowledge men's behavior as dangerous and take protective action for fear of appearing to be rude or even paranoid. Prevention programs concerned specifically with date or acquaintance rape are scarce, and the few studies available do not reveal a substantive impact on subsequent victimization rates compared with nonintervention controls.

2.3. Programs Directed at Mixed-Sex Audiences

In terms of their contents, programs directed at mixed-sexed target groups largely address the same aspects as single-sex programs, with a strong emphasis on providing information about the traumatic effects of rape

and on changing men's rape-supportive attitudes. Programs are also similar in terms of their short duration, rarely lasting more than 2 to 3 hours. If such programs are at all successful, their impact is short-lived and generally disappears within a few weeks. Although there may be practical advantages of combining male and female participants in a rape prevention program, this is outweighed by significant disadvantages. One is that men, particularly those with previous histories of sexual aggression, may be unwilling to approach their rape-supportive attitudes and coercive behaviors in an open way in the presence of women. In addition, the message—directed at female participants—that rape is prevalent and that the probability of prosecution for rapists is low may send the message to men that they can expect to get away with rape. Therefore, rape prevention experts recommend that rape prevention programs be directed at men and women separately.

2.4. Problems of Evaluation

Studies implementing and evaluating rape prevention programs are fraught with a number of conceptual and methodological problems that limit their significance. First, participation is voluntary, leading to self-selected samples of participants. In the case of men, this may lead to samples whose members are more open-minded and willing to learn and not representing men with strong rape-supportive attitudes and/or previous sexual aggression. In the case of women, self-selection may lead to the systematic underrepresentation of women who have previous histories of sexual victimization and who might find it too painful to approach the issue in a group setting. As a result, little is known about the success of rape prevention efforts in those target groups that are particularly in need of intervention. A second methodological problem concerns the operation of social desirability biases in the assessment of program success. Reports of greater victim empathy or lower support for rape myths may reflect respondents' tendencies to present themselves in a socially desirable way as much as they may indicate genuine change.

At the conceptual level, a major problem is the selection of criteria by which the success of an intervention is assessed. In the majority of programs, post-intervention measures of success are restricted to changes in cognition and affect such as reductions in rape-supportive attitudes and the promotion of victim empathy. As noted previously, significant but small changes are typically found on these variables immediately following the intervention. Far more important from an applied perspective is (a) whether cognitive and affective changes are of a sufficient magnitude and persistence over time to be able to exert an influence on behavior and (b) whether a program is successful in reducing the rates of sexual aggression. The evidence available suggests that the answer to the first question is no. Program effects on attitudes and victim empathy typically dissipate within 2 to 4 weeks after the intervention. The answer to the second question is hampered by the fact that there are few longitudinal studies that relate program participation to subsequent reports of sexual aggression in male participants or sexual victimization in female participants. However, the limited evidence that is available suggests that rape prevention programs so far have been largely unsuccessful in reducing the incidence of rape.

3. PREVENTION MEASURES AT THE SOCIETAL LEVEL

As a necessary complement to programs targeting individuals, measures in this category aim to change the sociocultural context in which sexual violence occurs. Their purpose is to increase the costs of rape for the perpetrators and to reduce opportunities for sexual assault in a society. Essential components of this approach are challenging gender stereotypes condoning aggressive behavior in men, reducing opportunities for sexual assault through policies designed to increase women's safety, and reforming the legal treatment of rape to raise the likelihood of conviction for rape offenders and improve the treatment of victims. Rape reflects a power differential between men and women that is part-and-parcel of a wider imbalance of power in society in Western culture and in most other parts of the world. Therefore, addressing this imbalance by promoting equal rights and equal participation of men and women in society will affect the perception of men's entitlement to women to gratify their sexual needs. In addition, gender stereotypes and sexual scripts placing men in the role of initiators and women in the role of gatekeepers of sexual interaction must be challenged and replaced by a more egalitarian view of how sexual intimacy should be negotiated.

Drawing on evidence that links sexual aggression to exposure to violent pornography, some have called for stricter control of accessibility of such material.

Moreover, education campaigns are needed to inform the public about the negative aspects of violent pornography with regard to promoting aggressive behavior.

3.1. Reducing Opportunities for Sexual Assault

This approach includes policy decisions designed to increase women's safety at the local community level. They include practical measures such as providing free late-night transport for women, designating safe spaces for women in public automobile parking lots, installing security cameras in public places, and ensuring that public pathways are well lit at night. These measures reduce opportunities for stranger sexual assault by making it harder for potential assailants to reach their targets without the risk of being caught and arrested. They can also serve to reduce women's fear of rape and to give women a greater sense of security. However, such measures are more likely to lead to a shift of sexual assaults to less well-protected locations than to a genuine reduction in sexual violence. In addition, their effectiveness is taken at face value and has not been subjected to systematic evaluation.

Different measures are required to reduce the incidence of date and acquaintance rape. Given a conclusive database that links both the perpetration of sexual assault and the risk of sexual victimization to alcohol consumption, legal regulations prohibiting the sale of alcohol to adolescents must be enforced. Because alcohol abuse on university campuses is widespread, particularly in fraternity cultures, special campaigns directed at curbing student drinking can also be instrumental in reducing sexual violence in this group.

3.2. Rape Prevention through Legal Reform

Among the most important objectives of rape prevention efforts at the sociostructural level is to increase the number of rapes reported to the police as well as the number of court cases that lead to convictions. There is no doubt that rape is a vastly underreported crime, as evidenced not least in the fact that victim surveys produce much higher incidence rates than do official crime reports. It is estimated that no more than 10 to 15% of all rapes are reported to the police. Victims' reluctance to report rapes has been attributed in large part to insensitive treatment of victims by the police. Many victims who report their rapes perceive police response as a "second assault," particularly if their cases depart from the real rape stereotype of a stranger assault. To encourage victims of rape to come forward, confidence must be established in the treatment that victims expect to receive from police, medical professionals, and legal institutions. To address this aim, police authorities have implemented procedures designed to ensure more sympathetic treatment of victims such as assigning female officers to the job of recording victims' statements. They have also developed training programs attacking derogatory attitudes toward rape victims and providing information about the traumatizing impact of rape. Despite these efforts, the reporting rate remains low compared with evidence based on large-scale crime surveys of the general population. Even if a rape is reported to the police, the likelihood of the case going to court is low and the probability of a conviction is even lower. Fewer than 10% of reported rapes end with convictions of the defendants. One reason for the low conviction rate in rape trials, particularly for date and acquaintance rape, lies in the tendency by defense lawyers to undermine victims' credibility by referring to the women's past sexual behavior or moral conduct. Rape shield legislation has been introduced to restrict the admissibility of past sexual history evidence. However, the law allows for exceptions to the rape shield provisions that open up the possibility of victim blame and reduce the likelihood of defendants being found guilty.

4. CONCLUSION

Rape prevention is a challenge that requires a multilevel approach, integrating measures directed at individuals as potential perpetrators and victims of sexual assault with macro-level approaches designed to change the sociocultural climate that is conducive to rape. The evidence reviewed in this article makes it clear that rape prevention efforts have had only limited success so far. Ultimately, a reduction in the prevalence of sexual aggression is the criterion by which the effectiveness of prevention efforts must be measured. To date, however, there is no substantive database that demonstrates the success of rape prevention programs in reducing incidence rates of sexual assault. For the time being, most prevention measures are judged on the basis of whether or not they are successful in changing the cognitive and affective antecedents of sexually aggressive behavior rather than in changing the behavior

itself. The restricted time span of most programs targeting individuals is undoubtedly a major limiting condition of their success. Given the deep-rooted attitudes and stereotypes underpinning sexual violence, it seems unrealistic to expect sustainable effects of an intervention lasting just a few hours. Moreover, the large-scale underreporting of sexual victimization is a major problem. So long as the true scale of the problem of sexual assault can neither be accurately measured nor reliably estimated, the ultimate success of rape reduction measures will remain hard to assess.

See Also the Following Articles

Aggression and Culture ■ Sexual Behavior ■ Sexual Behavior and Culture ■ Sexual Harassment

Further Reading

Abbey, A. A., McAuslan, P., Zawicki, T., Clinton, A. M., & Buck, P. O. (2001). Attitudinal, experiential, and situational predictors of sexual assault perpetration. *Journal of Interpersonal Violence*, *16*, 784–807.

Federal Bureau of Investigation. (2001). *Uniform Crime Reports*. www.fbi.gov.

Flores, S. A., & Hartlaub, M. G. (1998). Reducing rape-myth acceptance in male college students: A meta-analysis of intervention studies. *Journal of College Student Development*, *39*, 438–448.

Koss, M. P., Gidycz, C. A., & Wisniewski, N. (1987). The scope of rape: Incidence and prevalence of sexual aggression and victimization in a national sample of higher education students. *Journal of Consulting and Clinical Psychology*, *55*, 162–170.

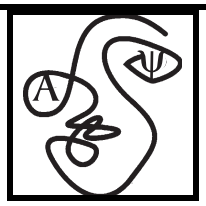
Krug, E. G., Dahlberg, L. L., Mercy, J. A., Zwi, A. B., & Lozano, R. (2002). *World report on violence and health*. Geneva, Switzerland: World Health Organization. www.who.int.

Rozee, P. D., & Koss, M. P. (2001). Rape: A century of resistance. *Psychology of Women Quarterly*, *25*, 295–311.

Schwartz, M. D., & DeKeseredy, W. S. (1997). *Sexual assault on the college campus: The role of male peer support*. Thousand Oaks, CA: Sage.

Spitzberg, B. H. (1999). An analysis of empirical estimates of sexual aggression victimization and perpetration. *Violence and Victims*, *14*, 241–260.

Yeater, E. A., & O'Donohue, W. T. (1999). Sexual assault prevention programs. *Clinical Psychology Review*, *19*, 739–771.



Reading Interventions

Laurice M. Joseph
The Ohio State University, Columbus, Ohio, USA

1. Introduction
2. Characteristics of Individuals with Reading Problems
3. Phonic Interventions
4. Whole Word Interventions
5. Repeated Readings
6. Vocabulary Development and Reading Comprehension Strategies
7. Principles of Teaching and Learning Reading
8. Targeting and Continuous Improvement of Reading Skills through Processes of Linking Assessment to Intervention
Further Reading

GLOSSARY

alphabetic principle Knowing that there are 26 letters in the standard English alphabet and that each letter is represented by a sound; knowing the symbols and corresponding sounds of one's own alphabetic collection (e.g. Cyrillic, Greek, and English).

morphemes The meaning units of words—roots, prefixes, and suffixes. With an understanding of morphological rules, a student may be able to find the root of an unfamiliar word and/or redeploy the term with alternative suffixes or prefixes.

orthography Conventionalized rules of letter patterns and spelling strategies (e.g., "I before E . . .").

phonemic awareness The alertness to and operation of discrete sounds in spoken language, including rhyme, rhyme production, sound blending, sound deletion and substitution, and sound segmentation.

phonological recoding The combining of phonological with orthographic knowledge about words. When a student is presented with a printed word, he or she may undergo a process of making letter–sound associations and matching the printed word to the spoken form of the word.

phonology Understanding the variety of possible sounds found in a language.

Reading interventions are necessary for children who struggle to acquire reading skills. There are several reading interventions for various types of reading problems, including word-level and comprehension problems. Reading interventions that are implemented should be tested on a continuous basis to determine their effectiveness and whether they need to be modified for individual children.

1. INTRODUCTION

Reading is a complex act. It is complex because written language is an abstract representational system of communication. Becoming literate is not a natural process but rather one that needs to be cultivated through various experiences. Children who are reared in families in which many rich verbal interactions occur are more likely to grasp the skills necessary to become literate. A rich language experience may include verbal dialogues between parents and children that consist of extending upon statements made by children. These types of dialogues facilitate vocabulary development as well as verbal reasoning skills. Before real reading takes

©2004 Elsevier Inc.
All rights reserved.

place, joint storybook reading with pictures provides children with rich language experiences. Parents can engage children in a discussion about the story by asking them to retell the story, explain why certain events might have occurred, and predict what might happen next as well as other types of comprehension questions. In 1991, Sulzby and Teal found that children progress from a spontaneous storybook telling phase (i.e., looking at pictures and telling a story about the pictures in an unconventional manner) to looking at pictures and telling a story about the pictures in a conventional manner as if they were reading the printed words. Finally, they actually read the printed words on a page without relying on the pictures. Rich language activities are necessary but not sufficient to becoming literate.

There are certainly other critical skills of emergent literacy, such as concepts about print, phonemic awareness, the alphabetic principle, orthographic processing, phonological recoding, morphological understandings, and reading fluency. Acquirement of these skills may serve to prevent future reading problems in children. A description of these critical reading skills is provided.

1.1. Concepts about Print

Before many children enter school, they begin to obtain concepts about print and this continues as children progress in school. Concepts about print involve skills in differentiating print from pictures, knowing that a book has a front and a back, and knowing that reading (in most Western languages) is a left-to-right and top-to-bottom process. At more advanced stages, children grasp that words written on a page can be spoken, and children know when a word and a sentence begin and end (i.e., spatial relationships between words and sentences).

1.2. Phonemic Awareness

Phonemic awareness is the alertness to and operation of discrete sounds in spoken language. It is one of the most important precursor skills to acquiring word recognition. Phonemic awareness activities include rhyme, rhyme production, sound blending, sound deletion and substitution, and sound segmentation. An example of a sound blending exercise is saying the individual sounds in a word such as /c-a-t/ and having children blend the sounds together to form a whole word. Sound segmentation involves saying a word as a whole, such as /ran/, and having children break the word apart and say each sound sequentially, such as /r-a-n/. These phonemic awareness activities should be explicitly taught to children in kindergarten and first grade according to a synthesis of research presented in the National Reading Panel Report. Among the phonemic awareness exercises, phonemic segmentation and blending were found to be the best predictors of word recognition performance in first grade.

1.3. Alphabetic Principal

In 1990, Adams asserted that children enter into the alphabetic system easily if they have acquired phonemic awareness. The alphabetic principle is knowing that there are 26 letters in the standard English alphabet and that each letter is represented by a sound. The use of alphabet books helps children grasp making letter–sound correspondences, particularly beginning letter–sound association of words.

1.4. Orthographic Processing

Orthographic processing consists of noting letter sequences in words or spelling patterns of words. Therefore, it involves examining the graphic structural properties of words.

1.5. Phonological Recoding

Phonological recoding involves combining phonological with orthographic knowledge about words. It is a process by which children examine graphic representations of a printed word and recode them back into its spoken language form. When children encounter a printed word, they go through a process of sounding out the word and then check to determine if the word matches a word stored in their memory and in their speaking vocabularies. Children initially make sequential one-to-one correspondences with letters and sounds. At more advanced stages in this process, they hierarchically decode words by using some letters in a word to cue the sounds of other letters (e.g., the e at the end of "came" cues the reader to say the /a/ as a long vowel sound).

1.6. Morphological Understandings

Conceptual knowledge about spoken words is critical for beginning readers. Morphemes represent the meaning units of words. Vocabulary development plays a

role in gaining morphological structures (meaning units) of spoken and written language, which is particularly critical when children, for instance, learn to read words that contain prefixes and suffixes as well as compound words.

1.7. Reading Fluency

The goal of basic reading instruction is to read words automatically or fluently. This is accomplished by acquiring a repertoire of word identification strategies as well as engaging in repeated readings of words in and out of context under timed conditions. Children are considered to be reading within grade level if they are reading between 75 and 100 words per minute correct on a grade-level passage with three to five errors.

2. CHARACTERISTICS OF INDIVIDUALS WITH READING PROBLEMS

Individuals with reading problems are generally characterized as having difficulties in obtaining word recognition skills and gaining meaning from text. These individuals are typically referred to psychologists for evaluation due to these general characteristics. Upon evaluation of several individuals with reading problems, the psychologist may discover that some children display delays across various areas, such as exhibiting low average cognitive ability (i.e., not in mental retardation range) and low average language, reading, math, and written language skills. These individuals are considered to be delayed types of readers, meaning that they tend to be delayed across cognitive and several academic learning domains.

Some individuals with reading problems have a central nervous system condition. These individuals are often classified as dyslexic and have average or higher intellectual abilities but severe reading problems coupled with speech production, rapid naming, phonological processing, and orthographic processing deficits. However, these individuals may not be delayed in math and other academic subject areas. Whether children are classified as dyslexic or the low average type, many children classified as having severe reading delays have deficits in phonological processing. Generally, students identified with learning disabilities in reading due to a central nervous system condition and a discrepancy between intellectual ability and achievement are not functionally different in their development of reading skills than those who are low average across several cognitive and academic domains. The diagnostic label is less important than the identification of the specific skills for which children experience difficulty. In fact, many researchers have found that low average readers and individuals diagnosed with central nervous system conditions or dyslexia exhibit similar underlying literacy deficits, particularly deficits in phonological awareness.

Other individuals' reading difficulties are due to poor instruction, which often leads to acquisition and fluency problems. Some students recognize and comprehend words accurately but not at a sufficient rate. Other students do not recognize and comprehend words accurately. The most efficient readers are those who recognize and comprehend words accurately and at a sufficient rate (i.e., reading 100 words per minute in the students' grade-level text).

3. PHONIC INTERVENTIONS

Phonics is the system by which children learn to make letter–sound correspondences. According to many prominent reading scholars, good phonics instruction does not dominate literacy instruction; it is brisk, lasting 20–30 minutes per day; and it includes phonemic awareness and spelling instruction. There are a variety of approaches to phonics instruction. Synthetic phonics instruction involves breaking words into their constituent sounds and blending the sounds together to form a whole word. An example of synthetic phonics instruction is Englemann and Bruner's Reading Mastery (formerly known as DISTAR). Analytic phonics involves the application of rules to identify words. Analogic phonics approaches involve word study methods of identifying unknown words by making analogies to known words. The research is inconclusive regarding the best approach to teaching phonics; however, some researchers suggest that synthetic approaches are more effective than analytic approaches. A synthesis of the research concluded that systematic direct instruction of phonics is more effective for achieving word recognition and spelling skills than indirect methods.

A number of contemporary phonic techniques involve studying phonological and orthographic features of words. Many of these approaches are not discussed here. However, some are provided as examples of techniques that can be employed to help children examine words. Word sorts are an anologic phonic approach for

helping children categorize words according to shared phonological, spelling, and meaning components. They can be in the form of closed sorts, in which the teacher establishes the categories, or open sorts, in which children induce the categories based on an examination of subsets of given words. Words are printed on index cards, and three or more words are designated as the category words. A stack of words is shuffled and then children are asked to sort the words into one of the categories according to common sound or spelling patterns. Bear *et al*. provide a comprehensive guide to creating various types of word sort lessons and other word study phonic activities. For instance, children can complete phonemic awareness sorts by sorting pictured objects according to beginning and ending sounds. This type of sorting has been found to be especially beneficial and appropriate for preschool and kindergarten children. Children can also spell words in respective categories as the teacher orally presents words. Howard Street and Early Steps, two effective tutoring programs for low achievers, incorporate word sort techniques rather extensively. Word sorts are also being used for children identified with mental retardation, and these exercises have been found to be helpful in improving these children's spelling performance.

Word boxes, a synthetic phonic approach, help children segment sounds of spoken language. This approach has been effective for helping children with learning disabilities identify and spell words and has helped first graders and preschool children achieve phoneme awareness skills. Typically, word boxes have been used in Clay's comprehensive reading recovery program. A word box consists of a drawn rectangle that has been divided into sections (boxes) according to individual phonemes in a word. Initially, children place tokens in respective sections as each sound in a word is articulated slowly. Eventually, children place letters (either magnetic or tile) in respective sections as each sound in a word is articulated. During advanced phases, children are asked to write letters in the respective divided sections of the box.

4. WHOLE WORD INTERVENTIONS

Whole word recognition or "look–say" approaches are effective especially for teaching words with irregular sound and spelling patterns. In look–say approaches, children practice reading words as a whole until they achieve fluency and the words become a part of their sight vocabularies. For instance, words that are printed on index cards can be read as a whole within a specified time period. Research has shown that whole word methods may produce immediate word recognition achievement effects, whereas phonic approaches may produce more long-term generalized effects, especially when students encounter words unknown to them.

5. REPEATED READINGS

Repeated readings are a very effective whole word recognition strategy for building fluency. Essentially, children read the same passage several times for 1 minute until mastery of the passage has been achieved. Mastery usually means reading the passage fluently—that is, reading all the words accurately and quickly. Reading fluency is highly correlated with reading comprehension.

6. VOCABULARY DEVELOPMENT AND READING COMPREHENSION STRATEGIES

According to McCormick, instruction aimed at developing vocabulary should consist of (i) teaching a small number of words intensively in any given lesson; (ii) exposing the student to the same word in various contexts and in a systematic and continuous fashion; (iii) including gamelike activities to stimulate interest in learning concepts, teaching morphological units (e.g., prefixes and suffixes); and (iv) explicitly, teaching derivations of words through exercises using the dictionary. Categorization activities help children group words that share similar meanings and help students make distinctions about words that are not similar. Cloze procedure exercises involving a passage with every "*n*th" word missing can be used to help children determine where words are placed contextually within a passage.

Learning strategies instruction approaches were found to be effective for helping children with learning problems attain concepts and improve their reading comprehension performance. Effective learning strategies instruction approaches consist of self-questioning, constructing mental representations to integrate information from text, and identifying text consistencies. Examples of these approaches include semantic mapping, PQ4R (preview, question, read, reflect, recite, and review), and reciprocal teaching.

Semantic mapping typically involves developing word webs that reflect students' understanding of

concepts and constructing a diagram connecting events of a story or connecting facts taken from content area textbooks. Students who struggle to grasp conceptual relationships may find diagrams to be helpful visual aids. Semantic maps can be either process oriented or product oriented. Process-oriented maps are usually completed before students read assigned material to help them establish some background knowledge. This type of mapping requires teacher facilitation of student responses. For example, a teacher may write a concept (e.g., satire) in a drawn box and ask the class members to give examples of when they experience it. As the class members respond, the teacher writes their comments below the concept. The teacher may then ask the class members how they feels when they witness the concept and to write their responses, drawing connecting lines to the other comments, and so on. Product-oriented maps, on the other hand, are produced by students as an outcome activity after they have read material. Students generally work independently or with other peers and construct a map connecting ideas presented in text.

The PQ4R method of comprehending reading material is an extension of SQ3R (survey, question, read, recite, and review). This method involves previewing the reading material, questioning the reading, reading to answer the questions, reflecting on the reading, reciting the reading, and reviewing the material. Previewing the material means surveying the chapter titles, main topics, and subheadings of the text. Students can turn the headings and subheadings into questions. Questions that were developed can be answered by reading the text. Reflections about the content can occur as the material is being read if the students pause to form connections and create images. Reciting is retelling what was read from memory. This form of retelling helps students monitor the information they are obtaining from the reading. If some information is not being retained, students may need to read sections of the text again. Lastly, students review the material by answering questions and referring back to the text for clarification of mistaken responses to questions. This method appears to be most appropriate for older students because it encompasses the application of higher level metacognitive processes. Although PQ4R is not a new method, it continues to be considered an effective technique for helping students organize and retain information from texts.

Reciprocal teaching is a reading comprehension approach that has helped delayed readers catch up to and even exceed the level of typically developing readers. This approach places heavy emphasis on teacher–student interactions in a rather cognitive apprenticeship fashion. After students and teacher read from common text, they dialogue with each other about the reading material. Initially, the teacher leads the dialogue by modeling strategies of predicting, question generating, summarizing, and clarifying text. The students are then asked to lead the discussions and apply the strategies that were demonstrated by the teacher. Guided practice is provided until students can use the strategies effectively. The goal of these reciprocal teaching interactions is to construct meaning from texts.

7. PRINCIPLES OF TEACHING AND LEARNING READING

The principles of teaching and learning word recognition and comprehension strategies include connecting to and building schema, demonstration and modeling, providing opportunities to learn and practice skills, and shaping reading behaviors through corrective feedback and reinforcement.

7.1. Connecting to and Building on Schema

Word recognition and comprehension skills are acquired more easily if the context from which these skills are presented is personally meaningful. This is especially important for children, who have limited conceptual knowledge. It is important for educators to become aware of children's prior knowledge and experiences so they can help children associate new information presented in text to their already existing schemas.

7.2. Shaping Reading Behaviors

Shaping behavior refers to modeling/demonstrating, providing corrective feedback, and eliciting reinforcers for successive approximations toward completing a task. Reading behavior is not acquired as naturally as language. In other words, we are not biologically predetermined to acquire reading skills but are biologically wired to acquire language. Therefore, reading behaviors need to be explicitly demonstrated and then modeled. It is important for parents or caregivers to provide opportunities for joint book reading early in children's development. Joint book reading activities should consist of the parent demonstrating reading

print on a page and looking at the pictures. When parents are not engaged in shared book reading with their children, they should model reading behaviors by reading, for example, a newspaper, magazine, letter, or novel. When children enter school, educators need to explicitly demonstrate prereading and reading skills, such as phonemic awareness, letter naming, and letter–sound correspondences, and whole word reading skills. They also need to model these skills while they are engaged in their own reading activities.

Children need corrective feedback on inaccurate responses made during reading activities. When children mispronounce a word, they should be provided with corrective feedback or prompted to supply the accurate pronunciation of the word. When children identify words incorrectly, they may be corrected through whole word correction procedures or phonic analysis procedures. Whole word correction procedures consist of saying a word correctly when children read a word incorrectly. Phonic analysis correction procedures may involve helping the child break the word apart by sounding out each letter or letter combinations and then blending the sounds to form a whole word.

Delivering reinforcers for efforts made toward achieving a goal can be considered as a way of supporting children through the reading process. Children with reading problems will struggle and become frustrated as they proceed through a trial-and-error process of learning how to read. Reinforcers need to be delivered for efforts made toward reading words accurately and comprehending text and not just for accurate performance. For many older children with reading problems, reinforcers may not have been a systematic part of the their instructional histories. In other words, consequences for reading behaviors may have not been emitted or delivered haphazardly rather than in successive approximations to desired reading performance.

7.3. Providing Opportunities to Learn

It is important for children to be provided with sufficient time to engage in reading activities. However, allotted time for reading is not sufficient to ensure active engagement in reading tasks. Instead, children need to be provided with many opportunities to be engaged in tasks during the allotted time for reading exercises. During oral and silent reading activities, children need plenty of opportunities to make responses. This can be achieved if lessons are structured so as to increase the number of responses children can make. For instance, choral reading of passages rather than round-robin reading increases all children's opportunities for being engaged in reading orally. Children also need many opportunities to practice skills that are learned. This is especially the case for children with learning disabilities and children with mental retardation. Overlearning leads to the ability to generalize skills to other contexts more easily.

7.4. Scaffolding

Scaffolding means providing necessary support and gradually reducing that support as an individual begins to function more independently. Educators need to employ scaffolds or necessary supports for children until they are able to complete tasks independently. Scaffolds include teacher facilitation and feedback as well as appropriate instructional materials that serve to make problems solvable. Examples of scaffolds used during reading activities are verbal prompts, context clues, and manipulative materials. Educators who use scaffolds often view themselves and the materials they design or select as mediators of the learner's development.

7.5. Motivating

Attribution retraining (training students to attribute their performance to the amount of effort rather than other factors) coupled with learning strategy instruction has been found to be effective for helping low achieving students to increase the amount of effort put forth on their reading through use of reading strategies. Children in kindergarten and first grade are more likely to expend effort toward reading even if they do not experience success, whereas older children, especially those with learning disabilities, are less likely to put forth effort if they do not experience immediate success with reading. This is because young children are likely to attribute their failures and successes to the amount of effort they expend, whereas older students are likely to attribute their successes and failures to other factors, such as teacher perceptions or task difficulty level. Likewise, many children lack an awareness of various strategies that would aid their reading skills, which makes it easy for these children to attribute their performance to other factors that may have no or a limited relationship to their performance.

8. TARGETING AND CONTINUOUS IMPROVEMENT OF READING SKILLS THROUGH PROCESSES OF LINKING ASSESSMENT TO INTERVENTION

The best approach to targeting specific reading skill problems is to conduct direct systematic assessment of reading performance. This can be accomplished through direct observations of students' oral reading behaviors in the classroom and through the use of curriculum-based reading measures. Curriculum-based reading measures consist of grade-level timed reading passages to assess fluency (accuracy and time) and also may consist of phonemic awareness, letter naming, and reading words out of context. Informal reading inventories and miscue analysis inventories are also helpful assessment measures for targeting specific reading performance. Evaluators should also observe the strategies students do and do not using while attempting to identify words and comprehend reading material. After specific skills and performance levels are assessed and noted, appropriate interventions can be targeted. Interventions should be implemented in a systematic manner and skills should be assessed during intervention periods to monitor student progress on a continuous basis as a function of the intervention. This process helps determine the effectiveness of the intervention. If a student does not demonstrate progress over time, then another intervention needs to be implemented and evaluated. This process of assessing and intervening is cyclical as students progress from performance level to another and from one skill to another.

See Also the Following Articles

Literacy, Improvement of ■ Reading, Teaching of ■ Teaching Effectiveness ■ Writing, Teaching of

Further Reading

Bear, D. R., Invernizzi, M., Templeton, S., & Johnston, F. (1996). *Words their way: Word study for phonics, vocabulary, and spelling instruction*. Englewood Cliffs, NJ: Prentice Hall.

Byrne, B., & Fielding-Barnsley, R. (1991). Evaluation of a program to teach phonemic awareness to young children. *Journal of Educational Psychology*, *83*, 451–455.

Clay, M. M. (1993). *Reading recovery: A guidebook for teachers in training*. Portsmouth, NH: Heinemann.

Connelly, V., Johnston, R., & Thompson, G. B. (2001). The effect of phonics instruction on the reading comprehension of beginning readers. *Reading and Writing: An Interdisciplinary Journal*, *14*, 423–451.

Daly, E. J., III, Witt, J. C., Martens, B. K., & Dool, E. (1997). A model for conducting functional analysis of academic performance problems. *School Psychology Review*, *26*, 554–574.

Ehri, L. C., Nunes, S. R., Stahl, S. A., & Willows, D. M. (2001). Systematic phonics instruction helps students learn to read: Evidence from the National Reading Panel's meta-analysis. *Review of Educational Research*, *71*(3), 393–447.

Engelmann, S., & Bruner, E. (1998). *Reading mastery I: Distar reading*. Chicago: Science Research Associates.

Gambebrell, L. B., Morrow, L. M., Neuman, S. B., & Pressley, M. (1999). *Best practices in literacy instruction*. New York: Guilford.

Joseph, L. M. (1998/1999). Word boxes help children with learning disabilities identify and spell words. *The Reading Teacher*, *52*(4), 348–356.

Joseph, L. M. (2000). Developing first grade children's phonemic awareness, word identification, and spelling skills: A comparison of two contemporary phonic approaches. *Reading Research and Instruction*, *39*(2), 160–169.

Joseph, L. M. (2002). Facilitating word recognition and spelling using word boxes and word sort phonic procedures. *School Psychology Review*, *31*, 122–129.

Joseph, L. M., & McCachran, M. (2003). Comparison of a word study phonics technique between students with moderate to mild mental retardation and struggling readers without disabilities. *Education and Training in Mental Retardation and Developmental Disabilities*, *38*.

Maslanka, P., & Joseph, L. M. (2002). A comparison of two phonological awareness techniques between samples of preschool children. *Reading Psychology: An International Quarterly*, *23*, 271–288.

McCormick, S. (1999). *Instructing students who have literacy problems* (3rd ed.). Englewood Cliffs, NJ: Prentice Hall.

Nation, K., & Hulme, C. (1997). Phonemic segmentation, not onset-rime segmentation, predicts early reading and spelling skills. *Reading Research Quarterly*, *32*, 154–167.

National Reading Panel. (2000). *Report of the National Reading Panel: Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups*. Rockville, MD: NICHD Clearinghouse.

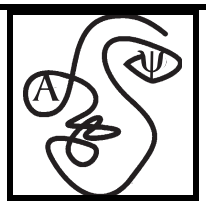
Nelson, K. (1996). *Language in cognitive development: The emergence of the mediated mind*. New York: Cambridge University Press.

Stahl, S. A., Duffy-Hester, A., & Stahl, K. A. (1998). Theory and research into practice: Everything you wanted to

know about phonics (but were afraid to ask). *Reading Research Quarterly*, *33*, 338–355.

Sulzby, E., & Teale, W. (1991). Emergent literacy. In R. Barr, M. L. Kamil, P. B. Mosenthal, & P. D. Pearson (Eds.), *Handbook of reading research* (Vol. 2, pp. 727–758). New York: Longman.

Vellutino, F. R., Scanlon, D. M., & Lyon, G. R. (2000). Differentiating between difficult to remediate and readily remediated poor readers: More evidence against the IQ-achievement discrepancy definition of reading disability. *Journal of Learning Disabilities*, *33*, 223–238.



Reading, Teaching of

Virginia W. Berninger
University of Washington, Seattle, Washington, USA

1. Historical Background
2. Relationship between Teaching and Learning
3. Developmental Framework for Creating a Functional Reading System
4. Learning to Read
5. Reading to Learn
6. Coordinating Reading and High-Level Skills
7. Individualizing Instruction for Students with Reading Disabilities
8. Optimizing Reading Instruction for All Students
Further Reading

GLOSSARY

alphabetic principle Set of correspondences between a phoneme and the alternative one- or two-letter spelling units that represent it in the writing system.

automaticity Nonstrategic, effortless, direct retrieval and production of a word.

decoding Converting letters to phonemes and synthesizing the phonemes into a pronounceable word or covert phonological representation of a word.

fluency Fast, smooth, coordinated oral reading that reflects the intonation (melody) of the spoken language or rapid, efficient silent reading.

morphemes Word parts that, alone or attached to other word parts, convey meaning.

morphological awareness Conscious reflection about and manipulation of morphemes, including prefixes, roots or bases (roots plus affixes), and suffixes (inflectional ones that mark number, tense, and comparison as well as derivational ones that mark grammatical speech part) in short-term and working memory.

nature–nurture interactions The notion that reading (and other) behavior is the result of biological and environmental influences.

orthographic awareness Consciously reflecting about and manipulating letter units in written words in short-term memory and working memory.

phoneme Smallest unit of sound that (a) has no meaning of its own but makes a difference in meaning and (b) corresponds to one or more spelling units.

phonological awareness Consciously reflecting about and manipulating sound patterns in spoken words (syllables, phonemes, and remaining parts of syllables after initial phonemes are removed) in short-term and working memory.

reading comprehension Text interpretation derived from vocabulary meaning, syntactic awareness, and discourse structures in text and from language and cognitive processes that go beyond what is stated in text.

rule-governed knowledge versus word-specific knowledge Abstracted patterns and relationships that can be applied across cases (single exemplars or multi-item constructions) versus learned representations in multiple formats (orthographic, phonological, and morphological codes) used in case (exemplar or instance) learning.

spelling unit or grapheme One or two letters that correspond to a phoneme.

strategies Nonautomatic reflective approaches applied to reading goals.

syntactic awareness Consciously reflecting about and manipulating the structures in sentences or conversational utterances in short-term and working memory.

©2004 Elsevier Inc.
All rights reserved.

Teaching reading is a teacher-led process of explicitly guiding students who actively participate in developing and coordinating (a) orthographic, phonological, and morphological awareness of words and their parts that results in accurate and automatic reading of words from the Anglo-Saxon, Latinate, and Greek layers of the language; (b) vocabulary meaning, syntactic awareness, and discourse structures that contribute to accurate and fluent reading and reading comprehension; and (c) interrelationships of reading with listening, speaking, and writing. Effective reading instruction also shows students how to apply the emergent knowledge and procedural skills to a wide variety of goals, including (a) learning from and thinking about written texts, (b) enjoying reading, and (c) using reading for pragmatic functions at school, home, and work as well as in everyday living.

1. HISTORICAL BACKGROUND

1.1. Writing Systems and Teaching Methods

Reading became possible when humans discovered three ways in which to represent the meaning or sounds of words in writing: ideograms/pictographs portraying concepts with line drawings, visual syllabaries standing for spoken syllables, and alphabets representing phonemes in spoken words with one- or two-letter spelling units. Three instructional methods corresponding to each of these writing systems were developed: whole word method (also referred to as look–say or sight words), syllable method and related word family method (syllable without its initial sound), and phonics (correspondences between phonemes and alphabet letters), respectively.

1.2. Debates

During the 20th century, scientific research resolved the debate about phonics versus the whole word. The conclusion was that phonics is the most effective, at least for the primary grades when basic word decoding skills are typically mastered. Research also showed that decodable text is effective in beginning reading instruction but that other methods, such as repeatedly reading the same whole words and teaching word families, are effective supplements to phonics. During the 20th century, scientific research also resolved another debate about whole language versus phonics that was about the optimal unit of instruction and instructional goal. Advocates of whole language argued for the whole text (based on good children's literature) as the unit and meaning as the goal, whereas proponents of the phonological method argued for the phoneme in spoken words as the unit and phonological decoding as the goal so that students can read independently. Scientific research supported the conclusions that (a) the phonological method is more effective for beginning reading instruction, (b) meaning is motivating to beginning readers who benefit from listening to and discussing good children's literature rather than reading it independently, (c) children exhibit individual differences in the amount and intensity of phonological instruction they require, (d) phonological instruction may be necessary but it is not sufficient, and (e) a balanced approach focused on both decoding and meaning is the most effective. Because language has multiple dimensions, effective instruction directs children's attention not only to phonology and orthography but also to the morphology, syntax, and discourse structure of spoken and written language.

The debate about direct instruction (which is teacher controlled with rote student responding) versus constructivism (which has minimal direct teacher input and relies on student construction of personal written language knowledge) is also being resolved with scientific research. Both teacher-led explicit instruction and strategy instruction for self-regulated learning benefit struggling readers, who learn from (a) direct explanations about the structure and rules underlying language and (b) teacher-led questions, prompts, or activities that encourage students to engage in active hypothesis testing in applying those explanations to their own reading. At the beginning of the 21st century, there is an increasing recognition that instructional methods for teaching reading should be grounded in scientific research rather than in philosophical belief.

2. RELATIONSHIP BETWEEN TEACHING AND LEARNING

Teaching has three instructional components: pedagogy (instructional behaviors of the teacher), curriculum (instructional goals and plans to reach those goals), and instructional materials (text and other instructional resources for achieving those goals). Teaching is related to learning but not in a simple one-to-one fashion. Teachers cannot program students' minds/brains through instruction the same way in which a

programmer can program a computer through input. Ultimately, students create their own learning (mental representations and organization) from the instructional cues they receive from pedagogy, curriculum, and instructional materials. Teacher knowledge of the developing brain, cognitive psychology, linguistics, and child development can influence the quality and appropriateness of the instructional cues provided. These instructional cues, in turn, are influenced by multiple biological (e.g., genes, neural architecture, maturation) and environmental (e.g., home literacy, school instruction, cultural, linguistic, socioeconomic) factors. Hence, learning is the by-product of nature–nurture interactions, and internal cognitive, affective, and motivational processes in the learner's mind/brain mediate the instructional cues and constrain but do not determine learning independent of environmental input.

3. DEVELOPMENTAL FRAMEWORK FOR CREATING A FUNCTIONAL READING SYSTEM

The functional reading system, which supports independent reading for a variety of purposes, is constructed over a long course of development from multiple language and nonlanguage systems in the brain and their interactions with the instructional environment. Although research has established correlations between beginning reading and subsequent reading skill, this relationship is not causal; developmentally and instructionally appropriate intervention at any time in reading development may improve the functioning of the reading system. In designing and implementing developmentally appropriate intervention, it is important not to confuse the novice and expert stages. Effective reading instruction aims just above the current instructional level to promote further growth along the developmental path to literacy. Children may follow different developmental paths to literacy. Some begin the journey at home during the preschool years, with parents using the lap method (reading storybooks to their children on a regular basis). Others do not begin the journey until school entry, but instructional methods have been developed to simulate the lap method in school settings (e.g., big books and little books with teacher-led finger point reading). However, federal programs are providing home literacy programs and incentives for explicit literacy instruction in preschools for those children who, heretofore, came to school unprepared for reading. Table I provides an overview of partially overlapping, cascading developmental stages and their instructional components.

TABLE I
Developmental Framework with Partially Overlapping, Cascading Stages

First stage: Learning to read words and understand text
1. Subword level: alphabetic principle, syllables, and morphemes
2. Word level: rule-governed decoding, word-specific representations, automatic word recognition
3. Text level: fluent oral reading, fluent silent reading, reading comprehension
4. Integration of reading with listening, speaking, and writing

Second stage: Reading to learn, enjoy, and work in school and nonschool settings
1. Domain-specific content areas in school curriculum
2. Leisure time activities in and out of school
3. Everyday reading for functioning in society
4. Vocation-specific reading for functioning at work

Third stage: Executive coordination of reading and high-level functions
1. Integration with technology (e.g., word processing systems, Web-based information systems)
2. Integration with thinking (e.g., critical thinking), mental world building (e.g., creative thinking), and problem solving

4. LEARNING TO READ

Basic decoding skills are typically taught during the first three grades. Fourth grade is an important transition time. Written text changes from mostly one- and two-syllable words of Anglo-Saxon origin that characterize primary grade reading material (and conversation) to a high density of words of more than two syllables and mostly of Latin (and Greek) origin that characterize reading material in the upper grades—when students continue to benefit from explicit instruction in decoding longer, more complex words. Also, the nature of the reading vocabulary, syntax in written text, and discourse structures changes. High-frequency vocabulary gives way to low-frequency complex vocabulary. Syntax in written material becomes more complex. Discourse structure changes from primarily narrative to primarily expository. Oral reading gives way to silent reading as the primary mode of instruction.

4.1. Teaching Word Reading Skills

Table II lists the instructional goals and associated instructional activities organized in the developmental progression from initial to mastery phases in learning word reading skills. In general, English word reading is predictable but not completely regular on the basis of alphabetic principle alone. Alphabetic principle can be applied to decoding both structure words (high-frequency words such as prepositions, conjunctions, pronouns, and articles that contribute to sentence meaning but have no meaning by themselves and are not always perfectly decodable) and content words (nouns, verbs, adjectives, and adverbs that tend to be more decodable). In English, a word might not be governed by regular rules of word identification at one level of analysis (e.g., single letter and associated phoneme) but might be when larger units (multiletter word families) or a different level of analysis (morphological) is taken into account. Although alphabetic principle benefits beginning word decoding, subsequent reading of polysyllabic words requires attention to additional linguistic cues (e.g., morphological) in word structure as well.

4.2. Teaching Reading Comprehension

In 2002, Carlisle and Rice provided an overview of teaching reading comprehension (see also Table III). Because lack of word automaticity (Table II) or text fluency (Table III) can interfere with normal reading comprehension, instructional activities aimed at developing oral and silent reading fluency are important components of reading comprehension instruction. From fourth grade onward, reading vocabulary instruction is necessary because many of the words are not part of students' speaking vocabulary as was the case with beginning reading texts. Effective vocabulary instruction requires far more than teaching the meaning of each word in the language. Both morphological awareness and syntactic awareness contribute to oral

TABLE II
Teaching Word Reading

Reading instructional goal	*Instructional activities*
Phonological awareness	Rhyming, syllable and phoneme analysis with spoken words
Orthographic awareness	Letter naming and letter writing; looking games requiring careful attention to (and memory for) all letters in a written word
Alphabetic principle	Producing the phoneme corresponding to a spelling unit containing one or two letters
Decoding monosyllabic Anglo-Saxon words (structure and content words)	Finding letters or spelling units that correspond to phonemes (applying alphabetic principle) and synthesizing the phonemes to pronounce the word
Decoding two-syllable Anglo-Saxon words (structure and content words)	Identifying type (e.g., closed, open, silent e, vowel team, r-controlled, -le) of each syllable in word and then decoding each syllable using alphabetic principle
Morphological awareness	Differentiating initial common spelling patterns that do or do not correspond to prefixes and final common spelling patterns that do or do not correspond to suffixes marking tense, number, or comparison or derivational suffixes marking grammatical part of speech; analyzing words into roots, affixes, or bases (affixed roots)
Decoding polysyllabic words from the Latinate and Greek layers of the language of Latinate and Greek origin	Introducing correspondences between spelling units and phonemes, syllable patterns (e.g., schwa), and morpheme patterns specific to the Latinate or Greek layers; practice applying multiple strategies based on alphabetic principle, syllable types, and morphemes to words
Automaticity	Practice reading monosyllabic and polysyllabic words from Anglo-Saxon, Latinate, and Greek word origin out of sentence context (on lists or flash cards) with and without rate criteria

TABLE III
Teaching Reading Comprehension

Reading instructional goal	*Instructional activities*
Oral reading fluency	Encourage children to attend to units larger than single words and to reflect the musical melody of the spoken language when reading aloud. Compute and record rate of accurate reading of repeated readings of same passages on which children can identify at least 90% of running words.
Silent reading fluency	Transition to silent prereading before oral reading, and follow same procedures as for oral reading fluency. Eventually transition to silent reading and wide reading of many passages. Record time for silent wide reading as a function of comprehension criteria (answering factual and inferential questions with 90% accuracy and story recall with main idea and supporting details).
Vocabulary	Guide students in inferring word meaning from sentence context. Define new vocabulary and use each word in a sentence to illustrate its meaning in context. Provide morphological patterns for generating new vocabulary items.
Sentence interpretation	Draw students' attention to the importance of word order by rearranging scrambled words into grammatically acceptable sentences. Discuss how meaning changes by reordering words slightly in a sentence. Help students paraphrase sentences to communicate the same idea.
Discourse understanding	Relate students' background (world) knowledge to content of the text to be read. Ask questions that prompt students to think about both information stated in text and information inferred from text. Teach strategies for self-regulating and monitoring the comprehension process, including prediction, summarization, clarification, and elaboration. Show how discourse structure can be used to build a mental representation of text. Describe the discourse structures of different genres: narrative (setting, characters, plot, and outcome) and expository (informational: lists, descriptions, comparisons, cause and effect, problem–solution; persuasive: one argument with evidence and conclusion with reasons, counterarguments with evidence, point–counterpoint, conclusions with reasons). Call attention to the different kinds of paragraph organization within the different genres (e.g., events for narrative, main idea and supporting detail for expository, mix of more than one genre).

vocabulary development as well as reading comprehension. Prior focus on question asking to assess comprehension has given way to pedagogy for explicit instruction in cognitive strategies for self-regulating and monitoring the reading comprehension process and for increasing active engagement of the reader in processing written text.

4.3. Integration of Reading with Other Language Skills

4.3.1. Listening

Students learn about reading by listening to the teacher's instructional language. Compared with conversational speech that is highly contextualized (embedded in social routines that support turn taking and communication repair), instructional language is decontextualized, with the focus redirected from language as a means of achieving an end (social communication) to language as an end in itself (language as the object of attention and reflection). Listening to the teacher read good children's literature stimulates linguistic development (vocabulary meaning, syntactic awareness, and discourse understanding) and cognitive development, and this in turn benefits reading development.

4.3.2. Speaking

Oral discussions about what has been read promote elaboration and deeper understanding of a text. Pedagogical methods for encouraging talk about reading material range from teacher-guided, collaborative, whole group discussions; to reciprocal teaching in which students assume teaching roles with some teacher guidance; to book clubs in which discussion is totally under student control.

4.3.3. Writing

Writing alphabet letters increases orthographic awareness, which in turn may facilitate learning to decode

words. Word spelling may exert reciprocal, mutually enhancing influences on the development of word reading. Written composition develops as children write about the material they read for various assignments.

5. READING TO LEARN

5.1. Teaching Domain-Specific Reading Skills in Content Areas

Although many middle school and high school teachers view their role as merely transmitting content matter, students learn more in content subjects when teachers also teach the vocabulary terms, textbook genre, and study strategies unique to the discipline.

5.2. Teaching Positive Affect Toward Reading

Unless children are successful in reading and enjoy reading, they are unlikely to engage in sufficient sustained practice of reading to become skilled readers. Therefore, reading instruction must create opportunities for children to learn to derive pleasure from reading. Some effective techniques for developing a reading habit and positive affect toward reading are including reading choice time at school and including leisure reading time at home. Children enjoy keeping reading logs of what they have read (e.g., books, magazines, newspapers) and competitions to see who can read the most books.

5.3. Teaching Everyday Reading Skills

Reading skills might not transfer automatically from the classroom to the real world without explicit instruction in specific reading tasks such as reading labels on food and household products, medicine labels and pharmacy instructions, menus, newspapers, magazines, advertisements, and the like.

5.4. Teaching Vocation-Specific Reading Skills

Vocation-specific and job setting-specific reading requirements vary considerably. High schools and colleges cannot prepare students fully for every kind of job-related reading goal (e.g., reading and applying information in highly technical manuals for operating heavy equipment or high-tech instruments). Job-related reading goals require not only decoding skills but also considerable domain-specific content knowledge. Some on-the-job reading comprehension strategies might be necessary.

6. COORDINATING READING AND HIGH-LEVEL SKILLS

6.1. Teaching Literacy for Technology Users

Although most elementary schools have computer laboratories, few have fully integrated technology into their literacy programs. By middle school, many students use word processing programs to complete written assignments based on reading source material but without explicit instruction in how to do so. As word processing and voice activation programs become more widely available and integrated with the instructional program, the need for explicit instruction in keyboarding, dictation, and integrated reading–writing systems for word processing will increase. During the information age, students also need explicit instruction in using Web-based information systems.

6.2. Teaching Critical and Creative Thinking and Problem Solving

Reading instruction should prompt readers to go beyond understanding a text at an initial level of analysis. For example, instruction may challenge readers to take a critical stance toward the text (e.g., its arguments, the author's intent), create new ideas that extend those in the text, and/or solve problems posed in writing assignments using reading source material. Many believe that teaching reading as a thinking skill is critical to the survival of democracy as a form of government.

7. INDIVIDUALIZING INSTRUCTION FOR STUDENTS WITH READING DISABILITIES

Effective strategies for teaching students with reading disabilities have been developed. However, many challenges remain in implementing individually tailored, appropriate instruction for all students with learning disabilities.

8. OPTIMIZING READING INSTRUCTION FOR ALL STUDENTS

At the beginning of the 21st century, expectations are rising for the educational achievement of all students. Considering the enormous normal variation among students at the same grade level, the expectation that all students will be readers and meet high-stakes standards will require reading instruction to become increasingly sophisticated.

See Also the Following Articles

Learning ■ Literacy, Improvement of ■ Reading Interventions ■ Writing, Teaching of

Further Reading

Adams, M. (1990). *Beginning to read: Thinking and learning about print*. Cambridge, MA: MIT Press.

Balmuth, M. (1982). *The roots of phonics: A historical introduction*. Baltimore, MD: York Press.

Berninger, V., & Richards, T. (2002). *Brain literacy for educators and psychologists*. San Diego: Academic Press.

Birsh, J. (Ed.). (1999). *Multisensory teaching of basic language skills*. Baltimore, MD: Paul H. Brookes.

Calfee, R., & Patrick, C. (1995). *Teach our children well: Bringing K–12 education into the 21st century*. Stanford, CA: Stanford Alumni Association.

Carlisle, J., & Rice, M. (2002). *Improving reading comprehension: Research-based principles and practices*. Baltimore, MD: York Press.

Chall, J. (1996). *Learning to read: The great debate* (3rd ed.). Fort Worth, TX: Harcourt Brace. (Originally published in 1967).

Clark, D., & Uhry, J. (1995). *Dyslexia: Theory and practice of remedial instruction* (2nd ed.). Baltimore, MD: York Press.

Ehri, L., Nunes, S., Stahl, S., & Willows, D. (2002). Systematic phonics instruction helps students learn to read: Evidence from the National Reading Panel's meta-analysis. *Review of Educational Research*, *71*, 393–447.

Nagy, W. (1988). *Teaching vocabulary to improve reading comprehension*. Newark, DE: International Reading Association.

Nagy, W., Osborn, J., Winsor, P., & O'Flahavan, J. (1994). Structural analysis: Some guidelines for instruction. In F. Lehr, & J. Osborn (Eds.), *Reading, language, and literacy* (pp. 45–58). Hillsdale, NJ: Lawrence Erlbaum.

Palincsar, A., & Brown, A. (1986). Interactive teaching to promote independent learning from text. *The Reading Teacher*, *39*, 771–777.

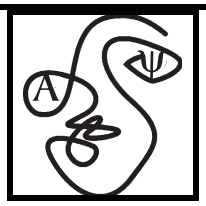
Rayner, K., Foorman, B., Perfetti, C., Pesetsky, D., & Seidenberg, M. (2001). How psychological science informs the teaching of reading. *Psychological Science in the Public Interest*, *2*, 31–74.

Swanson, H. L. (1999). *Interventions for students with learning disabilities: A meta-analysis of treatment outcomes*. New York: Guilford.

Thompson, G. B., & Nicholson, T. (Eds.). (1999). *Learning to read: Beyond phonics and whole language*. Newark, DE: International Reading Association.

Venezky, R. (1999). *The American way of spelling*. New York: Guilford.

Wong, B. (Ed.). (1998). *Learning about learning disabilities* (2nd ed.). San Diego: Academic Press.



Reciprocity Norm

Edwin P. Hollander

Baruch College and University Graduate Center, City University of New York, New York, New York, USA

1. Background
2. Mutual Interest
3. Interdependence
4. Social Exchange
5. Equity and Inequity
6. Bargaining and Negotiation
7. Implications
   Further Reading

GLOSSARY

adaptation level The standard of adjustment to a pattern of experience that a person has become used to.

comparison level A standard against which a person judges the attractiveness of a relationship with regard to whether he or she receives what is deserved, relative to his or her costs.

comparison level for alternatives The reference point employed in deciding whether to remain in an interaction. It can be considered as the lowest level of reward that an individual will accept in order to continue in the relationship. It also relates to the concept of adaptation level, with respect to what one has come to expect.

distributive justice A concern for whether an individual's investments in an ongoing relationship are or are not balanced by rewards.

egoism A self-centered basis for judging experience as good or bad in terms of one's benefits.

injustice A perceived negative outcome between actual rewards and expected rewards.

reference group A group that an individual takes as a standard for attitudes and actions, whether or not he or she belongs to it.

Social interaction between individuals involves an exchange of benefits, the most distinctive feature of which is the "give and take" of reciprocal behavior. This norm of reciprocity has been summarized by Gouldner as a universal norm stating that a benefit should be returned and, above all, that one who gives a benefit should not be harmed. These are often captured in everyday speech by such statements as "do unto others as you would have them do unto you" and "don't bite the hand that feeds you." Gouldner states that reciprocity fulfills two basic functions: to stabilize relationships and to serve as a check on egoism. These show concern for interdependence in relationships and the need to consider more than one's self-interest.

1. BACKGROUND

The rewards exchanged in relationships need not be of the same kind or of equal amount. Also, the anticipation of receiving a reward can account for such things as sacrifices, where goods are given to "appease the gods," which occur in many societies.

Another example is the distinction between a leader's election or appointment. An election has been found to give followers a great sense of responsibility and higher expectations for the leader's performance. In effect, the group gives the leader a benefit in advance by electing him or her to a position of higher status, and then the group feels a claim on the leader to return the benefit by producing favorable outcomes to the group.

©2004 Elsevier Inc.
All rights reserved.

2. MUTUAL INTEREST

When there is mutual interest, in most social settings one can assume that the other person understands the relevance of the reciprocity norm. This makes for a smoothing of relationships, if commonly understood. However, this need not always be the case, even with the need for interdependence. Most often, however, the concept of community of interest requires that reciprocity prevail, although it is conditional. There are variations among individuals and groups in the way in which it is applied. Although the broad culture may support the norm, its application still depends on the particular people in a given interaction.

3. INTERDEPENDENCE

Group support is closely tied to the motivation to take part in the task of a group. Whenever individuals come together to do things, they rely on one another for suitable performance. To the extent that members are rewarded by their attempts to achieve group goals, they will help work toward the goal and will be supported by the others. From a psychological standpoint, it is crucial that the individual care about what the group seeks to attain. Therefore, in joining groups, individuals are often attracted by the group's activity. Then they play a role that is rewarded by others' approval, which brings task and affiliation motives together.

Not just any source of support will influence a person's actions, however. The favorable effects of group support require two conditions: The individual must be positively motivated toward the activity, and he or she must be favorable about identification with the group members as relevant others. They must be people whose standards and support are desired—that is, a "reference group."

It is possible for an individual to do something he or she finds personally distasteful because others whose approval is sought are doing it. However, where a choice exists, the usual inclination is to be associated with those who support activities that the individual finds satisfying.

4. SOCIAL EXCHANGE

Although stated mainly in economic terms, social exchange goes beyond the simple notion of straight profit. Two early contributions to an understanding of social exchange derive from the 1961 work of Homans and the 1959 work of Thibaut and Kelley. Both involve certain concepts associated with a rewards–costs relationship.

Homans views social exchange as the underlying factor accounting for certain features of face-to-face interaction. He states that the rules for social interaction apply in various settings and can be understood primarily in terms of reinforcement, where the quantity and the value of an activity or a sentiment can affect the behavior of another individual. His propositions for what he calls elementary social behavior can be simply stated as follows: The more a person's activity is rewarded, the more likely the activity is performed; the more similar a present activity is to one rewarded in the past, the more it is performed; the greater the value of an activity, the more it is performed; the more rewarded an activity, the less any further unit of reward is valued; and the failure to receive an expected reward for an activity results in anger, which makes the results of aggressive behavior rewarding. In addition, he explains that a lack of reward leads to apathy, whereas too much of that reward decreases its value. Profit is defined as total rewards minus costs, with costs including values given up or forgone.

Thibaut and Kelley proposed that rewards for the parties involved are needed to maintain the interaction. In ongoing relationships, there is a trading off of rewards and costs for both parties. Thibaut and Kelley define rewards as

> The pleasures, satisfactions, and gratifications the person enjoys. The provision of a means whereby a drive is reduced or a need fulfilled constitutes a reward. We assume that the amount of reward provided by any such experience can be measured and that the reward values of different modalities of gratification are reducible to a single psychological scale. (p. 12)

They define costs as

> Any factors that operate to inhibit or deter the performance of a sequence of behavior. The greater the deterrence to performing a given act—the greater the inhibition the individual has to overcome—the greater the cost of the act. Thus cost is high when great physical or mental effort is required, when embarrassment or anxiety accompany the action, or when there are conflicting forces or competing response tendencies of any sort. (pp. 12–13)

In line with Homans' idea of profit as rewards minus costs, Thibaut and Kelley use the concept of comparison level. This is a standard against which the person judges

the attractiveness of the relationship with regard to whether he or she receives what is deserved, relative to his or her costs. Related to this concept is the comparison level for alternatives. It represents the reference point employed in deciding whether to remain in the interaction. It can be considered as the lowest level or reward that the individual will accept in order to continue in the relationship. It also relates to the concept of adaptation level, with respect to what one has come to expect.

The relationship of comparison level and the comparison level for alternatives can be stated simply: In general, an individual will be highly satisfied with outcomes if his or her comparison level and the comparison level for alternatives are both exceeded by actual outcomes. The opposite holds true with regard to dissatisfaction.

An injustice can also be thought of as an imbalance between an individual's rewards and costs in an activity. This is the fundamental point in Homans' concept of distributive justice. It deals with the condition in which an individual's investments are either balanced or not by his or her rewards. As previously noted, he considers profit to be rewards minus costs. Among these costs may be the investments of the individual—that is, what the individual brings to the situation from his or her background. This includes status, education, and experience, in addition to capability.

5. EQUITY AND INEQUITY

Equity is another term for fairness. An inequity is more likely to be perceived as an unfavorable comparison between one's own and others' rewards relative to costs. In the latter case, individuals will compare their rewards with others and take account of their own and the others' investments. In the workplace, these include such factors as seniority, responsibility, and productivity. Compensation is not the only element in assessing equity in the workplace. In 1983, Greenberg and Ornstein found that other variables, such as promotions, interesting work, pleasant coworkers, and increased status, contribute to the sense of equity. In 1987, Bandura more broadly asserted that goal accomplishment is a major feature of equity as well.

An earlier conception presented by Adams in 1965 defines equity as a function of the balance effected between a person's inputs and outcomes on the job. Inputs include qualifications and how hard one works at the job, whereas outcomes include pay, fringe benefits, status, and the intrinsic interest of the job.

$$\frac{\text{Outcomes}_{P}\ (\text{gains})}{\text{Inputs}_{P}\ (\text{costs})} \qquad \frac{\text{Outcomes}_{O}\ (\text{gains})}{\text{Inputs}_{O}\ (\text{costs})}$$

FIGURE 1 Comparison of a person's (P) own ratio of outcomes to inputs with an other's (O) ratio. Adapted from Adams (1965).

In his equity theory, Adams contends that the major expectancy is that an equitable input–outcome ratio will prevail. Inequity exists for a person (P) whenever he or she perceives that the ratio of his or her own outcomes to inputs and the ratio of other's (O) outcomes to inputs are unequal. This comparison of ratios can occur when P and O are in a direct exchange relationship or when both are in an exchange relationship with a third party, such as working in the same group or organization (Fig. 1).

In any consideration of justice or injustice, and equity or inequity, there is likely to be a difference of view. It is to be expected that P's perception of his or her own investments, costs, and rewards may not be the same as O's perceptions of them. However, it is not the actual fact nor the other person's perception that matter most in determining action. It is the individual's own definition of the situation. Therefore, inequities may be perceived by individuals despite the fact that others see no inequity in their version of social reality.

6. BARGAINING AND NEGOTIATION

Fundamentally, the process of interaction has a bargaining component. Even as they compete to gain desired outcomes, there are advantages to be achieved as P and O cooperate to trade rewards. Cooperation within competition is facilitated by negotiation. It remains the one method most widely used, and it is seen routinely in labor management relations, in which there may be awkwardness associated with power differences. To be effective, rules are necessary for an agreed procedural structure that depersonalizes the conflict. These help to avoid pitfalls such as perceiving the conflict as a zero-sum game, in which there must be a winner and a loser. Instead of accommodation, there may be an adherence to fixed positions, which may make both parties losers.

There is also the prospect of misunderstandings that derail progress. A basic feature of successful negotiation is to understand the opponent's position, which is helped by having open communication with the mediation of an outsider. A mediator can be useful for encouraging such things as limits on hostile language. This risk occurs when leaders desire to solidify their support base, with costs to the resolution of the conflict. Osgood's 1962 GRIT (Graduated Reciprocation in Tension-Reduction) grew out of his concern with the threat posed by the arms buildup in the Cold War between the United States and the Soviet Union. The strategy involves making concessions, usually small at first, to set in motion a process of reciprocation by the other side, ultimately leading to tension reduction of wider scope. At the negative extreme, in 1962 Tuchman described the misunderstandings and downward spiral that led to the arms buildup, always said to be "defensive," prior to the onset of World War I.

7. IMPLICATIONS

Running through these points is the norm of reciprocating, whether positively or negatively. The strategy utilized is dependent on the circumstances and the goals sought, including power relationships. If there are not mutual benefits in a solution, then the likelihood of a successful outcome is remote. Therefore, the norm is important, but it is sometimes not as important as the attitudes and motives of the parties involved.

See Also the Following Articles

Intergroup Relations and Culture

Further Reading

Adams, J. S. (1965). Inequity in social exchange. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 2). New York: Academic Press.

Bandura, A. (1987). *Social foundation of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice Hall.

Deutsch, M., & Coleman, P. (Eds.). (2000). *The handbook of conflict resolution: Theory and practice*. New York: Wiley.

Gouldner, A. W. (1960). The norm of reciprocity: A preliminary statement. *American Sociological Review*, *25*, 161–179.

Greenberg, J., & Ornstein, S. (1983). High status job title as compensation for underpayment: A test of equity theory. *Journal of Applied Psychology*, *68*(2), 285–297.

Hollander, E. P., & Julian, J. W. (1969). Contemporary trends in the analysis of leadership processes. *Psychological Bulletin*, *71*, 387–397.

Hollander, E. P., & Julian, J. W. (1970). Studies in leader legitimacy, influence, and innovation. In L. L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 5, pp. 33–69). New York: Academic Press.

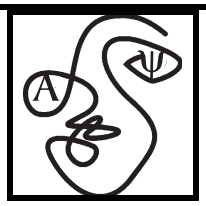
Homans, G. C. (1961). *Social behavior: Its elementary forms*. New York: Harcourt, Brace & World.

Osgood, C. E. (1962). *An alternative to war or surrender*. Urbana, IL: University of Illinois Press.

Pruitt, D. (1981). *Negotiation behavior*. New York: Academic Press.

Thibaut, J. W., & Kelley, H. H. (1959). *The social psychology of groups*. New York: Wiley.

Tuchman, B. (1962). *The guns of August*. New York: Macmillan.



Recruitment

James A. Breaugh
University of Missouri, St. Louis, Missouri, USA

1. Problems Resulting from Ineffective Employee Recruitment
2. Deciding on Recruitment Objectives
3. Recruitment Strategy Development
4. Considering the Job Applicant's Perspective
5. Assessing Recruitment Results
6. Conclusion
Further Reading

GLOSSARY

applicant population The type of job candidates (e.g., retirees and former employees) an organization hopes to reach by its recruiting efforts.

applicant self-insight Refers to the extent to which a job candidate has an accurate understanding of his or her abilities, interests, etc.

employee referrals A recruitment method that involves current employees of an organization referring people they know as candidates for open positions. Frequently, an employer will offer incentives to motivate such referrals.

realistic job preview (RJP) An attempt by an employer to convey accurate information concerning a position (e.g., job duties and work hours) to job candidates. RJPs can involve a variety of different approaches (e.g., a video, a brochure, and a visit to the work site) and are designed to improve the accuracy of recruits' job expectations.

recruiting methods Approaches such as job fairs and mailing lists used by employers to reach potential applicants for the purpose of informing them of job openings.

Research has shown that various recruitment decisions can have a significant impact on such factors as the diversity of job applicants and whether they accept job offers and remain on the job for a reasonable period of time. This article addresses such important topics as deciding on recruitment objectives, choosing a recruitment method, selecting recruiters, deciding on a recruitment message, and evaluating a recruitment campaign.

1. PROBLEMS RESULTING FROM INEFFECTIVE EMPLOYEE RECRUITMENT

The importance of effective employee recruitment has become increasingly recognized by organizations. This recognition at least partly stems from the fact that ineffective recruitment efforts (e.g., using unprofessional recruiters and presenting an overly positive view of an open position) have been linked to a number of serious problems that employers commonly face. For example, ineffective recruitment activities have been linked to (i) desirable job applicants turning down job offers; (ii) firms' Web sites being inundated

©2004 Elsevier Inc.
All rights reserved.

with resumes from unqualified job seekers; (iii) job advertisements resulting in applicant pools that do not reflect the demographic diversity desired; and (iv) newly hired workers, who had unrealistic job expectations, resigning after spending only a few weeks on the job. In order to understand how organizational recruitment activities are linked to these and other outcomes, several issues need to be considered, including what applicant population to target, how to reach these individuals, what recruitment message to convey, and whom to use as recruiters.

2. DECIDING ON RECRUITMENT OBJECTIVES

Prior to undertaking recruitment activities such as drafting a job advertisement or selecting individuals to serve as recruiters, it is important that an organization carefully considers its recruitment objectives. Although Breaugh reviewed numerous objectives that could be important to a given organization, this article focuses on four key questions that should be addressed in establishing recruitment objectives, and it shows how decisions concerning these objectives should guide subsequent recruitment actions. The first question an organization needs to answer is whether its focus is on filling a given position or on recruiting individuals who will progress through a series of positions over the course of a career with the organization. If the focus is on successfully filling a given position, then the employer does not need to be concerned with whether job applicants possess the talent and the motivation that would enable them to successfully perform higher level jobs.

A second important question is whether an employer prefers to fill jobs internally or externally. If the employer's focus is on recruiting from within its current workforce, it will not be difficult for the organization to reach the targeted population of potential recruits, nor is it likely that the firm will need to communicate information concerning organizationwide characteristics such as its values. In contrast, if an organization seeks to target applicants who do not currently work for it, more consideration will need to be given to how to reach desirable job candidates and what information to convey about the organization.

A third question meriting consideration in establishing recruitment objectives is whether an employer wants to attract a large number of applications or fewer applications from a select group of individuals who possess certain job-related attributes. If the former is the case, then posting a job advertisement that lists few applicant requirements for hiring on a well-known job board may be an appropriate strategy. In contrast, if an employer only wants to attract applications from a limited number of individuals who meet restrictive requirements such as expertise in SAS programming, then a different recruitment approach (e.g., a direct mail campaign) may be appropriate.

A fourth question that merits consideration is whether demographic diversity is important. To the extent that it is, an employer would be wise to incorporate certain steps into its recruitment plan. Such steps might include recruiting at colleges with diverse student bodies, having a diverse group of recruiters, and including pictures of minorities and women in recruitment brochures. Having given serious consideration to these four questions should enable an employer to more precisely define the applicant population it is interested in reaching.

3. RECRUITMENT STRATEGY DEVELOPMENT

Having established the types of individuals it seeks to attract (e.g., persons with experience working in a team or retirees who may be interested in working part-time), an organization needs to consider what method(s) to use to reach them. If an employer decides to target individuals who currently work for it, the organization may decide to post job announcements on the company's Web site. Alternatively, in order to limit the number of applications received, some employers rely on supervisors to nominate subordinates who are seen as both qualified and interested in a job opening.

To reach individuals who do not work for an organization, organizations have utilized several different recruitment methods, including newspaper advertisements, employment agencies, school placement offices, job boards, employee referrals, radio ads, direct mail campaigns, and job fairs. Although research has not been entirely consistent, research reviewed by Barber showed that employee referrals are frequently among the best recruitment methods. The benefits associated with employee referrals may be due to current employees (i) providing potential referrals with a realistic preview of what a job involves, (ii) prescreening potential referrals so that only individuals who are seen as

qualified and motivated are referred, and (iii) helping the referred person succeed once on the job. A potential drawback of using an employee referral program is that it may not result in a diverse pool of applicants if the current workforce lacks diversity.

Some employers are able to generate job candidates without any formal recruitment activities. For example, some individuals apply for jobs (i.e., mail a resume) without knowing whether a job opening exists. As noted by Breaugh and Starke, such direct applicants are more likely if an organization is highly visible and has a positive reputation. A number of highly visible organizations, such as Southwest Airlines, have relied heavily on their Web sites to generate job applications. Little research exists on whether Web sites generate more desirable applicants than more traditional recruitment methods.

A number of recruitment methods, such as job fairs and the use of college placement offices, involve the use of corporate recruiters. As discussed by Rynes and Cable, recruiters play an important role because they convey information concerning a job opening, they prescreen job candidates, and they answer questions from prospective employees. The selection and training of recruiters are also important because job candidates may interpret how recruiters treat them as a signal of how the organization treats its employees. Given the significance of their role, it is not surprising that considerable research on recruiters has been conducted. Although results are not consistent, in general, job candidates rate recruiters as being more effective to the extent that they are interpersonally skilled and informative. Research has shown that although training can have a modest effect on the performance of recruiters, the selection of recruiters is likely to have a greater impact. In particular, an employer should give consideration to whether a staff recruiter is used as opposed to using someone from the department with the job opening. The advantage of using a staff recruiter is that the person may be better able to answer questions about such issues as employee benefits and may be less likely to ask a candidate inappropriate questions. However, utilizing someone from the department in which the job opening exists (e.g., the supervisor to whom the person would report or a prospective coworker) offers important advantages. For example, in general, someone from the department will be seen as more informative concerning job duties, work group climate, etc. In addition, someone from the department, especially a potential coworker, is likely to be seen as a credible source of information. Given that many job seekers are cautious about accepting at face value all the information presented during the recruitment process, such credibility is important. Given their relative strengths and weaknesses, during the course of the recruitment process many employers use a team of recruiters composed of both line and staff recruiters. Regardless of the type of recruiter chosen, this decision should not be made lightly. Research by Rynes *et al.* showed that unprofessional treatment by a recruiter, such as failing to return phone calls or condescending behavior, can cause job candidates to withdraw as candidates.

After deciding on one or more recruitment methods to use to reach prospective employees, an organization needs to consider the information it wants to communicate to these individuals. In considering its message, one question an employer must address is whether it is interested in presenting a realistic view of a position or an overly positive one. A quick scan of the help wanted advertisements in a newspaper shows that many organizations convey an exaggerated view (e.g., can the pay for every job be competitive?). An employer should provide a realistic view of what a job involves. Presenting an exaggerated view of a position is unethical. Furthermore, it is likely to result in individuals accepting jobs that they quickly come to dislike. Wanous showed that such dissatisfaction is associated with a high rate of turnover for new employees. Such turnover can be quite expensive, especially if an employer has invested considerable training in the new employees or has reimbursed them for relocation expenses.

In phrasing a recruitment message, an employer should consider other factors in addition to realism. A key concern should be whether recruitment communications convey information that is considered important by recruits for making job choice decisions. Research has shown that most job candidates perceive information concerning the job (e.g., job duties and remuneration), the work group (e.g., coworker competence and whether interactions are cooperative or competitive), the supervisor (e.g., managerial style and technical competence), and the organization (e.g., corporate culture and economic conditions) as being particularly important. In addressing such topics, research has shown that recruits desire information that is sufficiently specific to allow for an informed job choice to be made. Thus, it is not sufficient for an organization to transmit information concerning how supervisors in general act or how coworkers in general relate. Instead, such information needs to be tied to a given job opening. Unfortunately, the recruitment messages sent by many organizations

fail to include the level of information specificity sought by many job candidates.

4. CONSIDERING THE JOB APPLICANT'S PERSPECTIVE

Filling an open position requires the agreement of both the hiring organization and the job candidate. Therefore, it is important that an organization consider the job applicant's perspective in planning the recruitment process. Because the types of information that many job applicants desire were discussed previously, it suffices to state here that an organization should do whatever it can to present such information in a credible manner. Off-the-record conversations with coworkers can be a particularly effective vehicle for conveying information that may be sensitive in nature (Does this organization truly have a merit-based reward system? Does the supervisor really involve employees in decision-making?). Such conversations are also effective for addressing the unique information needs of some job applicants, such as how easily a position will accommodate the observance of religious holidays.

If an organization seeks to hire new employees for whom there is good person–job fit, it is not sufficient for the organization to provide realistic and detailed information about important aspects of the position being filled. It also needs to be concerned with whether the job candidate can make sense of the information it conveys. For example, new employees may have been told that a job requires 60-hour workweeks or dealing with angry customers. However, if they lack relevant experience, some new hires may not be capable of fully grasping the meaning of the information conveyed by the organization. If an organization believes that inexperienced individuals will not be able to fully comprehend what a job involves, it may want to target individuals who can better appreciate what the job being filled demands. For example, individuals with prior experience working long hours should be better able to comprehend how a long workweek can interfere with family responsibilities. In order to maximize person–job fit, organizations should consider recruiting former employees whose prior performance with them was viewed favorably. If successfully recruited, former employees should have an excellent understanding of job and organizational attributes and thus have realistic job expectations.

5. ASSESSING RECRUITMENT RESULTS

After an organization has completed its recruitment activities for filling one or more job openings, it is important for it to evaluate the effectiveness of these activities. For example, an employer could compare recruiters, colleges visited, or advertising outlets (e.g., newspaper versus radio) in terms of the number of job applicants generated, the number of individuals hired, the cost per hire, the diversity of those hired, and/or the first-year performance of those hired. Unless an assessment of recruitment activities is conducted, an organization may not discover that certain activities have more value than others. For example, although employers frequently rely on job advertisements as a method for generating job candidates, research has shown that the use of employee referrals is a better method for generating employees who perform well and who remain on the job for a longer period of time. A particularly important source of information for assessing the effectiveness of recruitment actions can be newly hired employees and job applicants who were not hired. These individuals can provide insight into how they were treated during various stages of the recruitment process (e.g., whether phone calls were returned in a timely manner and whether a site visit was professionally conducted). Such information should provide those responsible for recruitment decision making with valuable ideas for improving the effectiveness of future recruitment efforts.

6. CONCLUSION

In staffing an organization, attention is frequently given to the soundness of a firm's selection system. Although using valid selection devices is clearly important, it is increasingly being recognized that poor recruitment activities can result in a pool of job applicants who lack the ability and/or motivation to successfully fill job openings. Such an applicant pool can restrict the value of even the best selection system. This article addressed a number of issues that can have a major impact on the overall success of a recruitment campaign. Specifically, it discussed key issues that an organization should consider in deciding whom to target for recruitment, how to reach them, what recruitment message to convey, and whom to use as recruiters. Thoughtful consideration

of these issues is likely to result in a more desirable applicant pool (e.g., more diverse, more qualified, and more interested) and in a more desirable group of new employees (e.g., higher job satisfaction and longer tenure).

See Also the Following Articles

Organizational Socialization ■ Person–Environment Fit ■ Personnel Psychology ■ Personnel Selection

Further Reading

Barber, A. E. (1998). *Recruiting employees*. Thousand Oaks, CA: Sage.

Breaugh, J. A. (1992). *Employee recruitment: Science and practice*. Boston: PWS-Kent.

Breaugh, J. A., & Starke, M. (2000). Research on employee recruitment: So many studies, so many remaining questions. *Journal of Management*, *26*, 405–434.

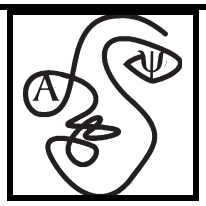
Greenhaus, J. H., Callanan, G. A., & Godshalk, V. M. (2000). *Career management*. Fort Worth, TX: Dryden Press.

Rynes, S. L., Bretz, R. D., & Gerhart, B. (1991). The importance of recruitment in job choice: A different way of looking. *Personnel Psychology*, *44*, 487–521.

Rynes, S. L., & Cable, D. M. (2003). Recruitment research in the twenty-first century. In W. C. Borman, D. R. Ilgen, & R. J. Klimoski (Eds.), *Handbook of psychology* (Vol. 12, pp. 55–76). Hoboken, NJ: Wiley.

Taylor, M. S., & Collins, C. J. (2000). Organizational recruitment: Enhancing the intersection of research and practice. In C. L. Cooper, & E. A. Locke (Eds.), *Industrial and organizational psychology: Linking theory with practice* (pp. 304–334). Oxford, UK: Blackwell.

Wanous, J. P. (1992). *Organizational entry*. Reading, MA: Addison-Wesley.



Rehabilitation Counseling

Timothy R. Elliott
University of Alabama, Birmingham, Alabama, USA

Doreen Miller
Southern University, Baton Rouge, Louisiana, USA

1. Introduction
2. Defining Roles and Functions
3. History and Evolution
4. Critical Legislation
5. Settings
6. Research Base
7. Summary
   Further Reading

Rehabilitation counseling is a counseling profession noted for its emphasis on serving persons with various physical, intellectual, behavioral, sensory, and developmental disabilities. The profession has been characterized by its counseling model and its responsiveness to state and federal mandates to coordinate the delivery of services to eligible individuals with disabling conditions to promote their employability and independence.

GLOSSARY

vocational evaluation A systematic and comprehensive assessment process to identify an individual's vocational potential; the evaluation may include, but is not limited to, self-report and observational assessment of a person's vocational and career interests, values, temperaments, work-related behaviors, aptitudes, skills, physical capacities, learning style, and training needs.

vocational rehabilitation The provision of services to enhance the employability and functional independence of a person who has been adversely impaired in these roles by a physical, intellectual, sensory, behavioral, intellectual, or developmental disability.

work adjustment The degree of congruence and correspondence between an individual's values and behavior and the behavior that is valued, accepted, and expected in a work environment.

1. INTRODUCTION

Rehabilitation counseling is a relatively young profession. In contrast to most human service professions, the birth and development of rehabilitation counseling are rooted in federal legislation. The genesis of the profession can be traced to the early 20th century in the United States. The establishment of the profession was precipitated by a number of social factors, the most fundamental of which was the need to provide rehabilitation services for an increasing number of people who were unable to work due to physical disabilities that were often incurred in the rapidly changing workplace. Although there was a need to provide services for these persons, there was no existing professional field available. Thus, the federal government mandated that the profession provide expert coordination and

©2004 Elsevier Inc.
All rights reserved.

administration of necessary services to citizens who were eligible for assistance.

2. DEFINING ROLES AND FUNCTIONS

Rehabilitation counseling initially focused on helping clients to become employable so that they could return to the workforce. This goal has expanded over the years to subsume a systematic process that assists persons with various disabilities to achieve their personal, career, and independent living goals in the most integrated settings possible. The rehabilitation counseling process involves communication, goal setting, and beneficial growth or change through self-advocacy, psychological, vocational, social, and behavioral interventions. The specific techniques and competencies used in rehabilitation counseling may include, but are not limited to, the following:

- Assessment, vocational evaluation, and appraisal
- Diagnosis and treatment planning
- Career (vocational) counseling
- Individual and group counseling treatment interventions focused on facilitating adjustment to the medical and psychosocial impacts of disability
- Case management, referral, and service coordination
- Program evaluation and research
- Interventions to remove environmental, employment, and attitudinal barriers as well as to enhance work adjustment
- Knowledge of disability-related legislation and regulations and of community resources
- Consultation services among multiple parties and regulatory systems
- Expert witness testimony
- Job analysis and modification, job development, and placement services, including assistance with employment and job accommodations
- The provision of consultation about access to relevant rehabilitation technologies

Currently, most rehabilitation counselors possess a master's degree in rehabilitation counseling (or in a related field) and are credentialed. The types of professional credentials available to a rehabilitation counselor fall into three general categories:

- Registration
- Licensure
- Certification

3. HISTORY AND EVOLUTION

Vocational rehabilitation (VR) emerged during the early years of the 20th century as federal policymakers wrestled with the consequences of an increasingly industrialized and urban economy as well as the ravages of international conflict. The Smith–Hughes Act of 1917 provided money to states on a matching basis to establish vocational education programs and also created the Federal Board of Vocational Education to administer VR programs. Relevant training and rehabilitation programs were provided under this legislation to meet current workforce demands. Parenthetically, this act established the first rehabilitation education programs within colleges of education in universities; rehabilitation counseling traditionally has been housed within this academic environment. The Soldier's Act of 1918 created VR programs for returning soldiers with disabilities, and the Federal Board of Vocational Education administered these programs.

Training programs were established in academic institutions, and considerable debate concerning the appropriate "intellectual home" for rehabilitation counseling ensued. One contingent viewed rehabilitation counseling within the larger domain of counseling psychology because rehabilitation counseling was perceived as a logical extension of the field's emphasis on vocational counseling, career development, and recognition of everyday people experiencing transitional stress in personal and social roles. Over time, however, critics expressed disapproval of "psychologized" coursework in graduate programs and argued that esoteric theories were of little use in the work setting. In addition, the multidisciplinary nature of VR, generally had little tolerance for theories that could not be easily conveyed or understood across disciplines. A greater emphasis on labor market issues, job assessment, work adjustment, and job placement was encouraged. The premium on practitioner competencies and issues permeates graduate training in rehabilitation counseling. Although coursework in areas considered "core" to psychology training may be offered, these are usually confined to areas of applied interest (e.g., assessment) and others may be permitted as electives (e.g., personality, abnormal behavior). In some institutions, rehabilitation counseling students may require the permission of the psychology department to enroll in any graduate psychology course.

4. CRITICAL LEGISLATION

Subsequent legislation increased the role of the federal government in preparing persons with disabilities for employment. For example, the Smith–Fess Act of 1920 increased the amount of money provided to state programs, and the 1935 Federal Social Security Act made the federal–state VR programs permanent. Other legislation extended VR services to blind individuals (the Randolph–Shepard Act of 1936) and to persons with mental retardation or psychiatric disabilities (the Barden–Lafollette Act of 1943).

The Vocational Rehabilitation Act of 1954 ushered in the "golden age" of rehabilitation and essentially created the rehabilitation counseling profession. It also dictated the model for systematically providing VR services and provided money to expand facilities. Services were also expanded to persons with substance abuse problems, public offenders, and the socially disadvantaged. The Rehabilitation Act of 1973 made services to persons with the most severe disabilities a top priority for state programs. It instituted the use of individualized written plans for clients and welcomed consumer input into the VR process. It also advanced a civil rights agenda for persons with disabilities by promoting access, prohibiting architectural barriers, and prohibiting discrimination in educational institutions and in the workplace.

Contemporary legislation has had a dramatic impact on rehabilitation counseling in terms of training and opportunities. The passage of the Americans with Disabilities Act in 1990 expanded the civil rights of persons with disabilities and reinforced the consultative role of rehabilitation counselors. However, a series of legislative acts appear to blur the support and role of rehabilitation counseling and its traditional role in the VR process. Specifically, the Workforce Investment Act in 1998 and the Ticket-to-Work/Work Incentive Improvement Act in 1999 were passed in part to circumvent perceived "red tape" that was associated with the traditional VR process as well as to encourage greater client input and options for rehabilitative services. These policies provide incentives for other entities in the private sector to offer and provide competitive rehabilitation services to consumers.

5. SETTINGS

From its inception, rehabilitation counseling has been virtually synonymous with state VR agencies, as originally mandated by federal legislation. This association continues to the current time. Traditionally, rehabilitation counselors have also found employment with Veterans Administration agencies, state employment agencies, state health departments, hospitals, mental health centers, prisons and correctional facilities, independent living centers, educational institutions (and training programs for students with disabilities), job training centers, social security offices, and workers' compensation programs. For decades, rehabilitation counselors were considered among the chief advocates for persons with disabilities and were seen as experts concerning these persons' rehabilitation and integration into society. Many rehabilitation counselors actively participated in the independent living movement for persons with disabilities.

However, changes in legislation and increases in the incidence of chronic disease and disability over the past three decades opened new opportunities in private business, substance abuse treatment programs, and employee assistance programs. Many rehabilitation counselors obtained positions in health care management programs, with insurance companies, and with third-party payers as case management specialists. Those in private practice often provide expert testimony, consultation, life planning services (for workers who incur lifelong severe disabilities), and vocational evaluation for legal firms, health care systems, businesses, and other agencies.

Individuals who possess a master's degree in rehabilitation counseling may pursue licensure as licensed professional counselors. With this licensure, they may be able to provide a wide range of mental health services for reimbursement. Most doctoral-level rehabilitation counselors retain their status as certified rehabilitation counselors. Others may qualify for psychological licensure in some states (depending on the graduate program content and the accreditation status of the department conferring the doctorate). These individuals may identify with rehabilitation psychology and pursue the diplomate in rehabilitation psychology offered under the auspices of the American Board of Professional Psychology. Doctoral-level rehabilitation counselors also assume academic positions where they are distinguished by their research, service, and training programs. Most rehabilitation counseling training programs are housed in colleges of education, a few are found in colleges associated with allied health professions, and some are found in psychology departments.

6. RESEARCH BASE

As a multidisciplinary enterprise that is constantly responsive and sensitive to the vagaries of federal legislation, fluctuating state budgets, and labor market trends, the VR literature has been criticized in academic circles for lacking theory-driven research (undermining its scholarly relevance to other fields) and an archival base (in which current knowledge builds on, advances, and refines prior knowledge). In a seminal text, Wright argued in 1980 that rehabilitation counseling has an established literature base that has steadily advanced the knowledge and sophistication of the profession. The peer-reviewed outlets associated with the field showcase a great diversity of topics, methodological rigor, and practitioner concerns. The *Journal of Rehabilitation* and the *Journal of Applied Rehabilitation Counseling* (both affiliated with the National Rehabilitation Association) have a clear emphasis on practitioner interests. The *Rehabilitation Counseling Bulletin* (affiliated with the American Rehabilitation Counseling Association) publishes studies of theoretical and practical interest and has a reputation for high standards in research design and methodology. The *Vocational Evaluation and Work Adjustment Bulletin* historically has focused on psychometric and practice issues germane to vocational assessment. *Rehabilitation Psychology* (affiliated with the American Psychological Association) publishes some work relevant to rehabilitation counseling. *Disability and Society* features work germane to rehabilitation counselors from a perspective associated with academic programs in disability studies and proponents of independent living programs for persons with disabilities. Other relevant outlets include the *Journal of Vocational Rehabilitation, Journal of Rehabilitation Administration,* and *Rehabilitation Education.*

The rehabilitation counseling literature has attended to the competencies and issues that face the field, including training, credentialing, and roles and functions. In addition, this literature has valued studies that address practice concerns associated with measurement, program evaluation, specific techniques and their application, and studies of psychosocial adjustment and vocational outcomes of clients with specific disability conditions. Rehabilitation counseling research often scrutinizes vocational evaluation practices and effectiveness as well as program and outcome evaluations of state VR programs. Recent research has attended to multicultural issues and practice implications of recent legislation. Many rehabilitation counselor researchers have considerable expertise in psychometric theory because the field has systematically evaluated the psychometric properties and appropriateness of psychological measures for use with persons with various physical, sensory, and intellectual disabilities.

A perusal of citation behaviors in available search engines suggests that research published in the outlets traditionally associated with rehabilitation counseling is somewhat insular because relatively few outlets outside of the field cite work published in these journals. A notable exception can be found in the contributions that rehabilitation counseling has made in the area of job placement. The practice of supported employment, which developed as a placement technique for persons with severe disabilities, has emerged as an effective placement strategy in several published clinical trials. The approach requires job specialists to observe and correct problematic client behaviors in the workplace over staggered and intermittent periods of time to promote work adjustment and model appropriate work behaviors. The practice has demonstrated efficacy in studies of persons with traumatic brain injuries, mental retardation, severe mental disorders, and sensory disabilities. Clinical trials supporting this technique have been published in mainstream (and influential) psychology and psychiatric journals.

7. SUMMARY

Rehabilitation counseling has evolved in response to federal legislation and labor market trends. Rehabilitation counselors in state agencies have been mandated to prioritize services for eligible clients with the most severe disabilities. During the time frame in which this mandate has been active, available data indicate that competitive employment placements increased substantially despite declining state budgets and increased expense for these clients. The profession will likely face competition from other professions that respond to recent legislation and provide VR services and placement to persons with disabilities, but it is probable that the field will continue its association with state agencies and thrive in case management roles in the private sector. Rehabilitation counseling has imbued its practitioners with unique skills in case management, job placement, and vocational evaluation,

and these skills serve the profession well in the public and private sector. Those with advanced training will also provide invaluable service in the program evaluation, training, and consultation that ensure the future of the profession.

See Also the Following Articles

Assessment and Evaluation, Overview ■ Job Analysis, Design, and Evaluation ■ Psychological Rehabilitation Therapies ■ Work Adjustment

Further Reading

Bond, G. R., Becker, D., Drake, R., Rapp, C., Meisler, N., Lehman, A., Bell, M., & Blyler, C. (2001). Implementing supported employment as an evidence-based practice. *Psychiatric Services*, *52*, 313–322.

Chan, F., & Leahy, M. (Eds.). (1999). *Health care and disability case management*. Lake Zurich, IL: Vocational Consultants Press.

Elliott, T., & Leung, P. (in press). Vocational rehabilitation: History and practice. In W. B. Walsh, & M. Savickas (Eds.), *Handbook of vocational psychology* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum.

Growick, B. (2000). The political implications of the TTW–WHA. *Rehabilitation Education*, *15*, 89–93.

McGowan, J., & Porter, T. (1967). *An introduction to the vocational rehabilitation process: A training manual*. Washington, DC: U.S. Department of Health, Education, and Welfare.

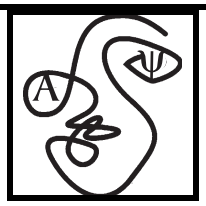
Parker, R. M., & Szymanski, E. (Eds.). (1998). *Rehabilitation counseling: Basics and beyond*. Austin, TX: Pro–Ed.

Rubin, S. E., & Roessler, R. T. (2001). *Foundations of the vocational rehabilitation process* (5th ed.). Austin, TX: Pro–Ed.

Shrey, D. E., & Lacerte, M. (1995). *Principles and practices of disability management in industry*. Winter Park, FL: GR Press.

Szymanski, E. M., & Parker, R. M. (Eds.). (1996). *Work and disability*. Austin, TX: Pro–Ed.

Wright, G. N. (1980). *Total rehabilitation*. Boston: Little, Brown.



Religion and Mental Health

Adam B. Cohen and Harold G. Koenig
Duke University Medical Center, Durham, North Carolina, USA

1. Introduction
2. Benefits to Mental Health
3. Applications of Religion
4. Caveats to Researchers
5. Testing Prayer?
6. Negative Implications of Religion
7. Training Counselors and Consultants
8. Conclusion
Further Reading

GLOSSARY

caregiver A person who takes care of patients with end-stage cancer or Alzheimer's disease; often a relative.

positive psychology Psychological research on positive emotions, such as joy, wonder, happiness, satisfaction, meaning, purpose, and hope.

In the past, the predominant view was that religion is inherently unhealthy. Recent research has begun to convincingly challenge this assumption. In hundreds of studies, religion has been associated with positive mental and physical health outcomes and with positive emotions. Although most studies have been cross-sectional, some have been longitudinal and several randomized clinical trials have also been completed. Despite these provocative associations, it is important to keep in mind several caveats. There may be significant scientific and theological concerns surrounding the prescription of religion and the empirical investigation of religion's effects on health. Moreover, not all religions may relate to health in the same ways, and in some cases religion may have decidedly negative consequences for health. Psychiatric and psychological training is beginning to emphasize a need to address religious and spiritual issues as part of clinical practice.

1. INTRODUCTION

For quite some time, many influential mental health professionals conceptualized religion and mental health as being diametrically opposed, and only within the past 10 years have such notions been challenged. Religion and spirituality have many components that are of potential relevance to mental health, including religious attendance, private religious activities (e.g., prayer and reading of holy scriptures), a feeling of connection or relationship with God or a higher power, religious beliefs, and religious coping. In addition, religion and spirituality may cause mental health via healthy lifestyles and behaviors and the promotion of social support. Some writers distinguish religion from spirituality by claiming religion is more organized, whereas spirituality has more to do with one's personal relationship with God or a higher power.

In 2001, Koenig *et al*. performed a comprehensive, systematic review of research on religion and mental health and identified 724 quantitative studies that

©2004 Elsevier Inc.
All rights reserved.

examined this association. Many different populations—from young to old, from sick to healthy—have been examined. Studies have been conducted in almost every area of the world, including the United States, Canada, England, continental Europe, Australia, China, Malaysia, Egypt, India, and Israel. The majority of studies (478; 66%) reported statistically significant relationships between religious involvement and better mental health, greater social support, or less substance abuse. Most of the studies were cross-sectional, but a number have been longitudinal and there are several clinical trials that have examined the effects of religious interventions on mental health. Longitudinal studies have also shown faster recovery from depression in community samples and more rapid adaptation to stressful life circumstances in religious caregivers. Several clinical trials examining the effects of a religious-based psychotherapy on treatment outcomes in depression, bereavement, and anxiety have likewise documented benefits for both Christian and Muslim patients.

2. BENEFITS TO MENTAL HEALTH

Mental health involves not only the freedom from mental disorders, such as depression, anxiety, psychotic conditions, or personality problems, but also the regular experience of positive emotions, such as joy, wonder, happiness, satisfaction, meaning, purpose, and hope. Relatively little research has been done on positive emotions and mental states. This is changing with the growth in "positive psychology," which has been led by Martin Seligman, former president of the American Psychological Association. Religiousness appears to be related to positive emotions, where the relationships may be even stronger and may have implications for the prevention of mental disorder. Of the approximately 100 studies conducted during the past century on religiousness and well-being, hope, and optimism, nearly 80% have found that religious persons experience significantly more positive emotions compared to those who are less religious. Similarly, in a recent study, measures of religious belief, coping, and spirituality were associated even more strongly than was social support with measures of life satisfaction among Protestants and Catholics, which is consistent with a role of religion-influenced positive emotion in well-being.

3. APPLICATIONS OF RELIGION

It is important to emphasize that clinical applications of religion must be theologically and ethically informed. For example, it is not only impractical but also may be theologically and ethically inadvisable to prescribe religious behaviors on the part of patients, for a clinician to impose his or her religious beliefs onto patients, or to create an atmosphere of intolerance for patients of diverse faith communities.

4. CAVEATS TO RESEARCHERS

From a research perspective, it is crucial to be theologically savvy. One aspect of this is to recognize that not all religions are the same, and that people from different faith communities may differ in psychologically relevant ways. One example is the salience of certain mental states (e.g., faith) to Jews and Protestants. For example, in the religion/well-being studies mentioned previously, social support from religious sources correlated modestly with well-being for Jews, Protestants, and Catholics. However, religious belief, spirituality, and religious coping were much more related to well-being among Catholics and Protestants than among Jews.

5. TESTING PRAYER?

Another aspect of being theologically informed in examining religion and mental health connections has been raised in the context of research on the effects on health of distant intercessory prayer. Chibnall *et al.* argued that such research may border on heresy because it tests or challenges God, that operational definitions of prayer may fly in the face of theological definitions, and that results are incredibly difficult to interpret. For example, do positive effects of prayer on health prove the existence of God? What do negative effects indicate? What if prayers by members of one faith community turn out to be more efficacious than prayers by those of another faith community?

6. NEGATIVE IMPLICATIONS OF RELIGION

Religious beliefs or activities may sometimes be associated with worse mental health or neurotic behavior.

Most clinicians today know from experience that religion may be utilized in an unhealthy manner or manipulated to serve defensive functions that ultimately impair mental health and prevent healthy growth. For example, religious concerns are common among people with obsessive–compulsive disorders and even people in nonclinical samples, including fears of sin, fears of God, and viewing the moral status of thoughts to be equal to that of actions. Vivid examples can be seen when individual cases are examined: the rigid fundamentalist; the obsessive–compulsive who attends church daily and prays 10 or 20 times per day; the psychotic patient who believes he is God, Jesus, or the devil; and the severely depressed person who is overwhelmed with guilt, believing that she has committed the unpardonable sin. This does not necessarily mean that religion promotes mental ill health, since ill people often turn to religion for comfort.

7. TRAINING COUNSELORS AND CONSULTANTS

Training programs in both psychiatry and psychology are beginning to emphasize the need for mental health professionals to address religious or spiritual issues as part of routine mental health care. This is now a requirement for accreditation of psychiatric residency programs by the American College for Graduate Medical Education (1994), which has mandated that training programs include didactic training in the "presentation of the biological, psychological, sociocultural, economic, ethnic, gender, religious/spiritual, sexual orientation and family factors that significantly influence physical and psychological development in infancy, childhood, adolescence, and adulthood" (pp. 11–12). Training programs in psychology are also required to ensure that the curriculum "includes exposure to theoretical and empirical knowledge bases relevant to the role of cultural and individual diversity," and religion is included in the definition of cultural and individual diversity.

As a result of increased training and exposure to research on the relationship between religion and health, there appears to be more widespread integration of religion into clinical practice, with therapists now utilizing patients' religious beliefs and social connections within the faith community to facilitate healing and recovery. Not only is the area of pastoral counseling growing rapidly but also there are indications that secular therapists are beginning to address spiritual issues when treating patients. Unfortunately, systematic research documenting when, where, and how frequent such integration is taking place has not been done.

8. CONCLUSION

Religion and mental health appear to be related—either for better or for worse. Clinicians need to be aware of the research that has been done in this area and understand how to both sensibly and sensitively apply this knowledge to clinical practice.

Acknowledgments

This work was partially supported by National Institute on Aging Grant AG00029 and a grant from the Positive Psychology Network to Adam Cohen. The views expressed do not necessarily represent those of these organizations.

See Also the Following Articles

Coping ■ Well-Being

Further Reading

Abramowitz, J., Huppert, J., Cohen, A. B., Cahill, S. P., & Tolin, D. F. (2002). Religious obsessions and compulsions in a non-clinical sample: The Penn Inventory of Scrupulosity. *Behaviour Research and Therapy*, *40*, 825–838.

Accreditation Council on Graduate Medical Education. (1994, March). *Special requirements for residency training in psychiatry*. Chicago: Accreditation Council on Graduate Medical Education.

American Psychological Association. (1995). Guidelines and principles for accreditation of programs in professional psychology. In *American Psychological Association Committee on Accreditation site visitor handbook* (pp. 5, I-27). Washington, DC: American Psychological Association.

Azhar, M. Z., & Varma, S. L. (1995a). Religious psychotherapy in depressive patients. *Psychotherapy & Psychosomatics*, *63*, 165–173.

Azhar, M. Z., & Varma, S. L. (1995b). Religious psychotherapy as management of bereavement. *Acta Psychiatrica Scandinavica*, *91*, 233–235.

Azhar, M. Z., Varma, S. L., & Dharap, A. S. (1994). Religious psychotherapy in anxiety disorder patients. *Acta Psychiatrica Scandinavica*, *90*, 1–3.

Braam, A. W., Beekman, A. T. F., Deeg, D. J. H., Smith, J. H., & van Tilburg, W. (1997). Religiosity as a protective or

prognostic factor of depression in later life; Results from the community survey in The Netherlands. *Acta Psychiatrica Scandinavia*, *96*, 199–205.

Chibnall, J. T., Jeral, J. M., & Cerullo, M. A. (2001). Experiments on distant intercessory prayer: God, science, and the lesson of Massah. *Archives of Internal Medicine*, *161*, 2529–2536.

Cohen, A. B. (2002). The importance of spirituality in well-being for Jews and Christians. *Journal of Happiness Studies*, *3*, 287–310.

Cohen, A. B., & Rozin, P. (2001). Religion and the morality of mentality. *Journal of Personality and Social Psychology*, *81*, 697–710.

Hill, P. C., Pargament, K. I., Hood, R. W., McCullough, M. E., Swyers, J. P., & Larson, D. (2000). Conceptualizing religion and spirituality: Points of commonality, points of departure. *Journal for the Theory of Social Behaviour*, *30*, 51–77.

Koenig, H. G. (2002). *Spirituality in patient care: Why, how, when, and what*. Radnor, PA: Templeton Foundation Press.

Koenig, H. G., McCullough, M., & Larson, D. B. (2001). *Handbook of religion and health: A century of research reviewed*. New York: Oxford University Press.

Levin, J. S., & Chatters, L. M. (1998). Research on religion and mental health: An overview of empirical findings and theoretical issues. In H. G. Koenig (Ed.), *Handbook of religion and mental health* (pp. 33–50). San Diego: Academic Press.

Pargament, K. I. (1997). *The psychology of religion and coping: Theory, research, practice*. New York: Guilford.

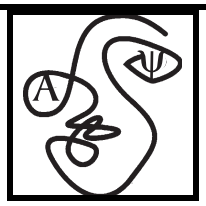
Propst, L. R., Ostrom, R., Watkins, P., Dean, T., & Mashburn, D. (1992). Comparative efficacy of religious and nonreligious cognitive–behavior therapy for the treatment of clinical depression in religious individuals. *Journal of Consulting and Clinical Psychology*, *60*, 94–103.

Seligman, M. E. P. (2002). *Authentic happiness: Using the new positive psychology to realize your potential for lasting fulfillment*. New York: Free Press.

Shafran, R., Thordarson, D. S., & Rachman, S. (1996). Thought–action fusion in obsessive compulsive disorder. *Journal of Anxiety Disorders*, *10*, 379–391.

van Oyen Witvliet, C., Ludwig, T. E., & Vander Laan, K. L. (2001). Granting forgiveness or harboring grudges: Implications for emotion, physiology, and health. *Psychological Science*, *12*, 117–123.

Zinnbauer, B. J., Pargament, K. I., & Scott, A. B. (1999). The emerging meanings of religiousness and spirituality: Problems and prospects. *Journal of Personality*, *67*(6), 889–919.



Residential Preferences and Attachment across the Lifespan

M. Vittoria Giuliani

Consiglio Nazionale delle Ricerche, Rome, Italy

1. Theories and Concepts
2. Housing Attributes and Preferences
3. Personal and Social Influences in Preference Ratings
4. Moving Home and Attachment
Further Reading

GLOSSARY

actual–aspirational gap theory A model that postulates that people perceive salient attributes of their physical environment and evaluate them based on certain standards of comparison, especially standards defined by what people believe they may reasonably aspire to.

appropriation of place The process by which people become more closely associated with or involved in places through taking possession, sharing and communicating social meanings, and manipulating and personalizing the environment.

cognitive dissonance A state of tension or discomfort resulting from inconsistency among beliefs or between one's beliefs and one's actions.

housing bundle The set of physical (e.g., dwelling size, building type, and neighborhood facilities) and nonphysical attributes (e.g., tenure) defining a housing alternative.

life values Desirable end states that the individual strives to attain in his or her life.

person–environment fit theory A model that assumes that congruence between personal preferences or needs and environmental press fosters environmental satisfaction and psychological well-being.

place attachment An affective bond between an individual and a particular place not interchangeable with another with the same functional quality.

place identity All aspects of self-identity having place-related implications.

proactive behavior Acting in advance to prevent potential stressors and to promote personal growth.

settlement identity Pattern of cognitions, feelings, behavioral tendencies, and skills that relate the identity of a person to a type of settlement and provide dispositions for future engagement with that type of settlement.

To what extent do places where people live meet their needs and desires? Family life and social life are changing dramatically in many areas of the world. How are these changes reflected in the meaning people attribute to their residence and the way they assess it? During the past few decades, researchers from different disciplines have addressed these issues from different perspectives, sharing the assumption that the quality of residential environment plays a central role in the global quality of life.

1. THEORIES AND CONCEPTS

What is the residential environment? As an environmental domain, the residential environment is usually

©2004 Elsevier Inc.
All rights reserved.

defined in contrast to other environmental domains, such as the working environment and the public environment. In both cases, the reference is to a concept of residence that is private, based on the model that has evolved in Western societies since the industrial revolution.

However, the home/work contrast has been questioned in the Western world both from a feminist perspective and in relation to the advent of telework, which has focused attention on a phenomenon—home-based work—that has in fact always existed but is currently growing.

Regarding the public/private contrast, cross-cultural research has revealed that the boundaries between the two spheres can be very fluid and do not necessarily coincide with the domestic confines. Institutional residences, such as student dormitories or nursing homes, can hardly be considered private environments.

Finally, it should be kept in mind that the residential environment has no definable spatial boundaries. One can approach residential preference at the level of room and dwelling unit or include the immediate surroundings, the neighborhood, or even the larger community. The term housing bundle, used to indicate the pattern of attributes defining a housing alternative, includes items pertaining to different scale units (i.e., from both the micro- and the macroenvironment). Research has shown that for different groups of people, salient environmental units may be different; furthermore, when assessing housing bundles, including both the dwelling unit and the neighborhood dimension, people modify the importance assigned to the attributes of housing or neighborhood when evaluated in isolation.

This article focuses on the domestic environment. Aspects related to neighborhood satisfaction and preferences are reviewed elsewhere in this encyclopedia and will only be mentioned here in reference to the location of the dwelling unit.

1.1. The Meaning of Residential Environment and Home Attachment

In the study of residential environment, a distinction is regularly made between house and home. A house (or apartment) is defined by its physical structure, and a home is defined by a set of meanings. Research on housing is focused on the spatial and behavioral aspects of living, whereas that on the home deals with the symbolic and emotional meaning of the living experience. The term dwelling has been put forward as a link between house and home: Dwelling means the process by which a physical entity takes on psychological, symbolic, and social meaning, transforming it into a home.

Many psychological and sociocultural dimensions have been identified to distinguish between home and nonhome and to grasp personal differences in the home meaning: security and control, reflection of one's ideas and values, acting upon and modifying, permanence and continuity, place for relationships with family and friends, center of activity, refuge from the outside world, indicator of personal status, material or physical structure, and a place to own.

A comparison of home and nonhome reveals that the meanings attributed to the former generally have a positive connotation, whereas those associated with the latter are generally negative in physical as well as social or personal terms. Recently, there has been an increasing focus on the negative aspects of home experience, viewed, for example, as a constraint (for the old or the unemployed) or as a place of possible violence or oppression, particularly for women and children. In addition, extending analysis of the meaning of home to different cultures, other meanings emerge, such as religious meanings as evidenced, for example, by the beliefs of Hindu and Muslim women.

To characterize the variety of representations and use of space in homes in different cultures, and also the changes that occur in the same culture over time, Altman and colleagues adopted a dialectic approach in which the sociopsychological meaning of the home is given by a limited number of bipolar dimensions that characterize the individual's relationship with society, such as identity/communality and accessibility/inaccessibility (or openness/closedness). In relation to preferences, it can therefore be said that cultural and individual differences are represented by the changing balance in the relative strength of the oppositional poles that define each dimension.

Many researchers have argued that what defines the very nature and essence of home, as distinguished from house, is the emotional relationship with the place. Based on earlier theoretical work on place and home by phenomenologists and humanistic geographers, a growing body of psychological literature has been devoted to the study of affective bonds that people develop with places. The concept used in this context is place attachment. Although it is argued that people may develop feelings of attachment to places of different spatial scale—residence, neighborhood, city,

region, country, etc.—and to a variety of types of setting, from the work environment to natural settings, empirical research in this area has focused on residential settings. Although from a spatial standpoint environmental units are "nested" within each other, to be attached to one's city does not necessarily imply being attached to one's neighborhood, house, or apartment: The degree of attachment to different environmental units varies in relation to sociocultural and demographic variables, such as social position, age, and gender.

Attachment has been considered in relation to residential satisfaction—even though the direction of the relationship is not clear (e.g., is one more satisfied because one is more attached, or is one more attached because one is more satisfied?)—to residential mobility, and to residential choices. With respect to satisfaction, attachment represents a comprehensive measure, including even behavioral and emotional aspects that go beyond a mere affective response. In fact, what qualifies attachment is not the positive valence of affects but the fact that it is perceived as a bond, with an enduring quality, directed toward a specific target, not interchangeable with any other with the same functional quality. Consequences of attachment have been postulated at the level of social involvement and mobility and at the level of well-being and mental health.

Moreover, attachment is not necessarily positive; it can represent a source of individual hardship or be the result of conformist attitudes or social rejection. It has also been suggested that place attachment develops and supports place identity; that even in mobile countries such as the United States, the majority of people form affective bonds with places; and that changes of place consistent with changes occurring in their needs and identity during their lifetime may actually favor rather than hinder the formation of affective bonds.

1.2. Preferences, Satisfaction, and Housing Choices

What kind of residential environment is good for what kind of person? Environmental psychologists have tried to answer this fundamental question by studying residential satisfaction and preferences. Although the terms satisfaction and preferences are often used interchangeably to indicate what people like, they are two distinct concepts. Both have to do with an assessment, but in the former case the fulfillment of needs is evaluated, whereas in the latter case the fulfillment of desires is evaluated. Furthermore, although satisfaction is generally measured in relation to the actual environment, preferences can also refer to an ideal environment (i.e., not bound by the limitations of the individual or the constraints of the market). Finally, strictly speaking, preferences imply a choice among alternatives.

Preferences and satisfaction have been put in direct relation in different models of residential satisfaction, such as the actual–aspirational gap theory or the person–environment fit theory. Both models postulate that satisfaction derives from the perceived congruence between personal preferences or aspirations (which are in turn influenced by the personal characteristics of residents) and the objective and subjective characteristics of a place. Satisfaction has therefore often been used as an indirect indicator of preference. Moreover, both preferences and satisfaction have been considered to be predictive of behavioral intentions and actual behavior. However, various studies have shown a low correlation both between choices and preferences and between satisfaction and preferences. Although the explanation in the first case refers especially to external factors (e.g., market restrictions) and only to a lesser degree to other factors, such as the negotiating process with other members of the household, the discrepancy between satisfaction and preference is generally explained by psychological factors, such as processes of adaptation and assimilation or mechanisms of cognitive dissonance. Highly positive responses often reflect or even conceal an adjustment to inevitable inadequacies.

How is the meaning attributed to the environment related to the preferences toward the objective attributes of the housing? The subject has not been investigated systematically. Most research on housing preferences has simply been aimed at identifying correlations between objective environmental attributes and subjective evaluation (satisfaction and preferences) and/or personal and social influences on subjective evaluation. The two lines of research on subjective evaluation and the psychological meaning of the residential environment have largely remained separate.

A more general question is whether residential preferences and choices can be predicted from subjective beliefs about the effect of different housing alternatives on the attainment of life values (e.g., comfort, family, freedom, security, togetherness, and money). Unfortunately, limited attention has been devoted to this issue. Results provide evidence that beliefs determine the evaluations. In addition, they suggest

important differences between residential preferences and simulated choices, particularly that preferences do not completely explain choices and that choices, unlike preferences, are better explained by the overall evaluation of all attributes than the evaluation of single attributes.

1.3. Evaluative Perspectives and Aesthetic Preferences

Asking someone to express his or her preference for a residential environment is like asking whether that person would like to live in a certain place. However, preferences can be different depending on whether the place would be a permanent residence or a vacation home, a place to live alone or with the family, etc. Furthermore, as already mentioned, a residence has to satisfy a multiplicity of needs, not the least of which is aesthetic. Satisfaction of aesthetic needs is a salient component of housing satisfaction with regard to the inside, the outside, and the neighborhood, and the aesthetic qualities of the living environment are given high importance across all age and social groups. However, empirical research on preferences for interior decoration and furnishing has shown that aesthetic preferences can enter into conflict with other needs, such as comfort, practicality, or upkeep, and that some persons will change their preferences depending on whether they are asked to evaluate a house for living purposes or merely from an aesthetic standpoint.

The explanatory models adopted in environmental psychology are classified into two different traditions: (i) the experimental and motivational aesthetics of Berlyne, who traces aesthetic preferences to the physiological arousal response elicited by the "collative properties" of the objects, which include structural properties (e.g., complexity and coherence) and aspects of experience (e.g., novelty and the ability to surprise), and (ii) the sociopsychological tradition that derives preferences from the social meaning of the objects (i.e., their ability to express social identity in terms of belonging and distinction). Other theories have emphasized the cognitive components of aesthetic preference: For example, according to the schema theory, preferences relate to the discrepancy between the perceived example and the abstract schema, or prototype, representing the appropriate category. Increases in discrepancy evoke arousal and interest, but preference is highest for moderate to low discrepancies. Considerable empirical work has been carried out to study the physical and symbolic properties of a building and what individual variables affect preferences.

1.4. Methods for Assessing Preferences

In studying preferences, a distinction widely made is that between revealed and expressed preferences: Revealed preferences are inferred from actual behavior, whereas expressed preferences are directly derived from the answers people give to questions about what they like or dislike. Although revealed preferences are used considerably in the sociological and economic domains, in psychology the second technique is preferred. Persons may be asked to provide preferential judgments that record their personal likes or dislikes concerning the places to be assessed (e.g., through adjective checklists and rating scales), or they may be asked to record comparative appraisals (i.e., evaluations made with regard to some implicit or explicit standard of comparison). Preferences can also be surveyed indirectly by measuring residential satisfaction, although attention must be paid to possible discrepancies between the two measures, as previously mentioned.

One useful approach is based on various forms of trade-off games that allow for priorities for competing interests to be identified. For example, answers to the question of how persons would change their home if they had one room more or less provide a good ranking of priorities.

Variety in methods and techniques is also determined by the presentation of the stimulus to be evaluated: verbal descriptions, drawings, photographs, scale or life-size simulations, or direct observation. There is a lack of data concerning the relative validity of different methods and measurement procedures, even though it has been shown that observers' perceptions of room size and crowding vary significantly in three different modes of environmental display (actual room, video tape, and scale model).

Finally, there is a major methodological difference between asking for the evaluation of separate housing attributes and asking for a global evaluation of different housing alternatives. Adoption of one or the other method is of considerable theoretical importance. In the former case, it is assumed that the preference for a particular housing alternative is determined by the sum of the evaluations of its attributes, whereas in the latter case emphasis is placed on the holistic

dimension of the evaluation, hence the preference for a single attribute depends on the other attributes with which it is associated.

2. HOUSING ATTRIBUTES AND PREFERENCES

What are the objective qualities that make one residence preferable over another? The question has been dealt with in a variety of ways, by a number of disciplines, and in the theoretical and applied spheres, ranging from market research to public agencies responsible for housing policies. Considered here are selected aspects that best exemplify the psychological approach at three environmental levels: domestic interior, architectural style, and housing location.

2.1. Interior Arrangement and Decoration

A house becomes a home through a process of "appropriation," which largely consists of choosing and changing the functional organization of space, furniture, and decoration—in other words, by personalizing it. A place that can be personalized becomes a symbolic mirror of certain aspects of the personality of the person residing in that place. It has been shown that people can derive hypotheses regarding the social and personal aspects of the inhabitants from photographs of their living rooms, and that observers' inferences on the personality and lifestyle of the homeowners correspond significantly to their self-concept.

In a survey of all members of 80 families in Chicago, valued and cherished objects in the home were found to have different meanings based on age and gender. Women and older adults tended to value them as symbols of personal relationships and family ties. Younger people and men, on the other hand, were more utilitarian.

Studies carried out in Italy and France to identify the descriptive dimensions of the physical environment that best differentiate interior decorating choices found that in addition to the kind and style of furniture (modern vs old-fashioned, rich vs plain, etc.), the arrangement of the furniture plays an important role. The items characterized by particular complexity and richness (e.g., antique and large or highly decorated furniture) are generally associated with a symmetrical or central position, whereas more sober items are associated with a more elaborate and unconventional composition. The close correspondence between preferences for interior design and furnishings and cultural background and social status suggest that individual taste is largely shaped by cultural ideologies.

If interior decorating choices largely reflect the sociocultural context, this holds true even more for the internal organization of space, over which individuals can exercise less control. Anthropological studies have shown that division of the domestic space and the functional restrictions by gender, age, or activity are congruent with the social and political organization of a culture. In addition, architectural and organizational characteristics have a precise psychological meaning. For example, by asking participants to construct models of houses that would promote in turn privacy, security, and social interaction for the occupants, researchers found that privacy houses had significantly more rooms and corridors, security houses were smaller, and social interaction houses had greater visibility between rooms.

Environmental psychology has emphasized societal trends related to socialization patterns and household roles that are reflected in the actual configuration of the house (room size and functional organization) and in changes in needs and desires. For example, a study of prototypical house plans in the United States in the second half of the 20th century found that increasing the size of the kitchen and incorporating it more openly into the rest of the house seems to respond to women's desire for greater sharing of domestic tasks and greater interaction among family members. More informality in social interaction is reflected in the widespread preference for multifunctional spaces, though with considerable variations related to household composition (e.g., the presence and age of children) and life stage.

2.2. Architectural Style

Research on aesthetic evaluations of buildings is both of theoretical interest, because it allows for the verification of theories on aesthetic perception in a real context, and of applied interest in relation to urban design policies. Various studies in North America, Australia, the United Kingdom, and Turkey have investigated the preferences for different house styles. As with interiors, at least with regard to single houses, preferences for the exterior were largely influenced by the symbolic meaning transmitted by the architectural form. Affective and social meanings are attributed to the different styles: A particular style may be considered more pleasing or interesting than

another, and inferences can be drawn about the characteristics of the residents.

The architectural traits associated with social status are resistant to change and learned at an early age. By looking at photographs of homes built at the end of the 19th century in Boston, observers were able to identify the occupational status and social class of the original owners; schoolchildren (6–12 years) showed high internal agreement in the cost ranking of several house styles, with preferences being correlated with cost.

In addition to confirming the U-shaped relationship between complexity and liking theorized by Berlyne, of particular interest is the lack of correlation found in many studies, if not all, between local experience and preferences. In evaluating representative examples of Australian and British houses, students in Glasgow and Sydney, for example, liked the ones to which they were less accustomed.

Using a set of stimuli including drawings of six different house styles (contemporary, colonial, farm, Tudor, saltbox, and Mediterranean), ample convergence of preferences of laypersons for the Tudor and farm styles was found in different geographic areas of the United States. Differences were found, using the same and/or other stimuli, between adults and children, men and women, architects or students of architecture, and people without a background in design.

2.3. Housing Location

An important characteristic of the place of residence is the location. Residential location preferences have been approached from two perspectives: (i) the importance attributed to a variety of services and functions (e.g., the proximity of parks, schools, friends, and relatives) that characterize a specific location and (ii) the typology of the setting, particularly in relation to the distance from city center. The two aspects are not independent, but since the evaluation of neighborhood attributes has been studied mainly in relation to neighborhood satisfaction, attention is concentrated here on the second aspect.

A wide literature, especially in North America, has repeatedly noted the strong preferences for small towns or suburban living that accompany the desire for a single-family, self-owned home. If the individuals' age and role in the family are taken into consideration, the picture becomes much more complex. Preferences for one location or another are determined by familiarity and expectations, largely guided by cultural norms, about the proper type of place for a person to live at a particular stage in life. The psychological ties established with certain kinds of settings have been conceptualized in terms of settlement identity. Settlement identify is formed through personal experience of specific types of residential environments and the generalization of the characteristics to prototypes of places. Thus, city center is associated with greater opportunities for leisure and cultural activities, a wider range of contacts, and greater accessibility to certain services. The suburbs are associated with more quiet, better air quality, greater homogeneity among residents, etc. Small towns seem to be associated with the best characteristics of both the city center and the suburbs. Identification with one kind of setting or another and the consonance or dissonance between settlement identity and actual place of residence influence both preferences and assessment of different settings.

3. PERSONAL AND SOCIAL INFLUENCES IN PREFERENCE RATINGS

As mentioned previously, in addition to environmental characteristics, a wide variety of individual characteristics, ranging from personality and personal history to demographic and sociocultural factors, influence the preferences and meanings attributed to a house. Here, attention is focused on only a few characteristics, keeping in mind that the separation between them is rather artificial and that differences arise mainly through the interaction between different dimensions.

3.1. Designers vs Users

Numerous studies have documented that architects or other design experts assess built environments differently from laypeople and that architects are unable to predict lay evaluations. Regarding housing style, there is a higher consensus in preferences among trained people than among untrained laypersons, with architects preferring more "high style," "atypical" houses. It has been suggested that this is the result of a greater predisposition among architects for higher levels of sensory input and their familiarity with atypical examples acquired during training, which leads to the formation of prototypical schemes that are different from those of untrained persons.

3.2. Gender Role

In all societies, men and women have different roles in relation to the environment and the opportunities offered by it. This is reflected in both the significance attributed to the home and the needs and preferences related to the residential environment at various levels. Although there is a high degree of consensus between men and women regarding the essential qualities of a home, there are differences in the salience of different attributes: For many women, the home symbolizes self, family, and social relationships; for many men, it symbolizes ownership and childhood memories. Women attribute greater value to objects of "contemplation" (photographs, plants, etc.); men attribute greater value to objects symbolizing action, such as sports equipment and stereos. Regarding housing design, modern style is preferred by males, whereas women tend to favor more traditional design.

Women are generally less satisfied with housing than men and desire more privacy. However, this difference in satisfaction with a living situation often does not depend on a difference in assessment criteria but, rather, on the subjective fact that many women, particularly single mothers or older women living on their own, are more poorly housed than men in analogous situations.

Women, particularly single mothers, are more open to alternative housing arrangements, such as cohousing, that allow them to share some domestic tasks. Other differences are found at the neighborhood level: Proximity to friends and accessibility to services are more important for women, who are also more in favor than men of denser communities and multifamily housing.

3.3. Age and Stage of Life

Some of the most important transitions in people's lives are either defined by a change in where they live (e.g., leaving the parental home) or constitute a reason for moving (e.g., the birth of a child or an older child leaving the home). Furthermore, changes in the significance attributed to the residential environment and the needs it has to satisfy are associated with different stages in life.

Use of the spaces making up the residential environment and their salience at the psychological level vary during the course of a lifetime. For children, interest is centered on play: The home has to be safe and provide space inside and outside for play. As adolescence approaches, privacy and self-expression take on a preponderant role. Older teenagers and young adults living in their parents' home show little interest for the characteristics of the house, as long as privacy and control over their private space are ensured. Loss of control seems to characterize places that young adults experience as the most depressing. Families with children are often dissatisfied with the space in the home and want to move to residential neighborhoods far from the city center, whereas many older couples and single residents prefer to live closer to the city center.

Housing for the elderly has increasingly become a key issue in environmental psychology. However, placing an undue emphasis on the concept of the elderly as a homogeneous group is a mistake. First, old age encompasses an arc of life that, with the increase in life expectancy and advances in health care, is becoming elongated and ever more differentiated—to the point that distinctions are now made between young old, old, and old old. Second, interindividual differences tend to accentuate with age in relation to the variety and quality of life experience. Finally, cultural differences are heightened: For example, whereas independent living is a desire shared by the vast majority of the elderly in the United States, in Japan the solution preferred by an elderly widow is to live with her son.

Some features nevertheless seem to characterize the relationship between the house and advancing age: most notably, a strong attachment to the house, seen as a symbol of independence and a material trace of the pleasant and unpleasant events of one's own life. Numerous studies have shown the negative consequences on health from moving from one house to another and, even more, from a house to an institution; however, studies have also demonstrated the moderating effects of the circumstances in which that change takes place, particularly whether the choice is voluntary, its preparation, the characteristics of the new environment, and the social context before and after. Increasing emphasis is being placed on the proactive aspects of housing choices in old age in an attempt to identify the strategies for optimization activated by a change in environment.

4. MOVING HOME AND ATTACHMENT

Prolonged association between an individual and a place is widely recognized as one of the distinctive features of

attachment to place. As happens with attachment to people, individuals may not be conscious of their attachment to a place and may only become aware of it in particular circumstances, such as when the bond is threatened. Moving means breaking this bond; therefore, relocation has been considered as providing unique insight into the attachment processes.

The early theories on attachment take two opposing approaches. The first approach stresses the emotional experience of places: Affective ties are established automatically and unconsciously with one's place of residence and are destroyed by mobility. The second approach focuses on cognitive rather than emotional aspects of people–environment relationships and stresses the positive function of change. Affective bonds with place are the result of a positive assessment of the congruence between the attributes of the environment and the inhabitant's needs. Because needs are liable to change during one's lifetime, if one's needs are fully satisfied by a new place, attachment can be expected to develop even after a move.

Recent research, while acknowledging the importance of emotional aspects, emphasizes the dynamic relationship between person and place and the modification of affective bonds throughout a person's lifespan. However, still lacking are longitudinal studies on how new bonds are formed, to what extent one place of attachment can be replaced by another, or whether multiple attachments can coexist.

Although leaving a place one loves is almost always stressful and painful, not all people appear to develop feelings of attachment with their place of residence and the reaction to a change is closely linked to a persons' time and place orientations (i.e., their perception of themselves as anchored more to a past or to a future timeframe).

Finally, two factors of fundamental importance are the possibility of remaining in contact with the places of the past and the perception of control with respect to the change. Enforced relocations, caused by natural and technological disasters, urban renewal, or transfer into institutions, constitute a break in the individual's and community's sense of continuity that is often underestimated and cannot be compensated by the physical quality of the new environment.

See Also the Following Articles

Environmental Design and Planning, Public Participation in ■ Environmental Stress ■ Person–Environment Fit ■ Residential Satisfaction and Perceived Urban Quality ■ Urban Environments and Human Behavior

Further Reading

Altman, I., & Low, S. M. (Eds.). (1992). *Place attachment. Human behavior and environment. Advances in theory and research* (Vol. 12). New York: Plenum.

Altman, I., & Werner, C. M. (Eds.). (1985). *Home environment. Human behavior and environment. Advances in theory and research* (Vol. 8). New York: Plenum.

Aragonés, J. I., Francescato, G., & Garling, T. (Eds.). (2001). *Residential environments: Choice, satisfaction and behavior*. Westport, CT: Bergin & Garvey.

Arias, E. G. (Ed.). (1993). *The meaning and use of housing: International perspectives, approaches and their applications*. Aldershot, UK: Avebury.

Bonnes, M., Giuliani, M. V., & Werner, C. M. (Eds.). (1987). *Home interior:* A European perspective [Special issue]. *Environment and Behavior*, *12*(2).

Cooper Marcus, C. (1995). *House as a mirror of self*. Berkeley, CA: Conary Press.

Craik, K. H., & Zube, E. H. (Eds.). (1976). *Perceiving environmental quality*. New York: Plenum.

Csikszentmihalyi, M., & Rochberg-Halton, E. (1981). *The meaning of things. Domestic symbols and the self*. Cambridge, UK: Cambridge University Press.

Giuliani, M. V. (2003). Theory of attachment and place attachment. In M. Bonnes, T. Lee, & M. Bonaiuto (Eds.), *Psychological theories for environmental issues* (pp. 137–170). Aldershot, UK: Ashgate.

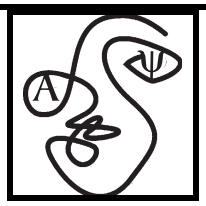
Lawrence, R. J. (1987). *Housing, dwellings, and homes: Design theory, research and practice*. Chichester, UK: Wiley.

Lawrence, R. J. (Ed.). (1991). The meaning and use of home [Special issue]. *Journal of Architectural and Planning Research*, *8*(2).

Nasar, J. L. (Ed.). (1992). *Environmental aesthetics: Theory, research, and applications*. Cambridge, UK: Cambridge University Press.

Rapoport, A. (1990). *The meaning of the built environment: A nonverbal communication approach* (R. ed.). Tucson, AZ: University of Arizona Press.

Tognoli, J. (1987). Residential environments. In D. Stokols & I. Altman (Eds.), *Handbook of environmental psychology* (pp. 655–690). New York: Wiley.



Residential Satisfaction and Perceived Urban Quality

Marino Bonaiuto

University of Rome La Sapienza, Rome, Italy

1. Introduction
2. Urban Environment Quality in Experts' and Laypersons' Evaluations
3. Describing Residential Satisfaction and Perceived Residential Environment Quality
4. Explaining Residential Satisfaction
5. Applications
   Further Reading

GLOSSARY

(perceived) environmental quality indexes [(P)EQIs] Standard set of (perceived) indicators for evaluating a specific place (e.g., residential PEQIs) that can also be used for policy and monitoring functions.

place attachment Forms and degrees of affect (e.g., feelings, moods, emotions) that people develop over time and come to experience with reference to the places in which they are born, live, and act.

(residential) postoccupancy evaluation (POE) An examination of the effectiveness for human users of occupied designed (residential) environments.

residential satisfaction The experience of pleasure or gratification deriving from living in a specific place, that is, the global evaluations that inhabitants give of their housing that can be considered at various levels of scale (e.g., house, building, neighborhood).

residential (urban) quality Evaluation of the various specific attributes of the residential (urban) environment; this evaluation can be technical and expert based (i.e., "objective") or observer based (i.e., perceived or "subjective").

The quality of the urban environment, and of the residential environment in particular, is one of the main components of people's overall quality of life. Global satisfaction with people's own residential environment has both affective and cognitive components. On the one hand, people experience feelings with respect to their own houses/neighborhoods (i.e., place attachment); on the other hand, they assess the degree of quality for each different feature of their own houses/neighborhoods (i.e., residential quality). Knowing residential satisfaction, residential attachment, and residential quality with reference to the inhabitants of a specific place can help its design, management, and policy.

1. INTRODUCTION

The quality of the urban environment, and of the residential environment in particular, is one of the main components of people's overall quality of life, although it tends to rank lower (in terms of importance in the overall quality of life) than do items such as leisure activities, economic conditions, work, friendships, and marriage/family life.

©2004 Elsevier Inc.
All rights reserved.

Residential satisfaction can be defined as the experience of pleasure or gratification deriving from living in a specific place, that is, the global evaluation that inhabitants give of their housing that can be considered at various levels of scale (e.g., house, building, neighborhood). Conceptually, it includes the three main components of the psychological construct of attitude: cognition, affect, and behavior. Operationally, residential satisfaction is often measured through inhabitants' answers (ranging from "not at all" to "completely") to items such as "How satisfied are you with the house/building/neighborhood?" Satisfaction scores lack absolute validity; a high score cannot be automatically considered as an indicator of optimal performance of a housing program. However, satisfaction scores do have relative validity; other things being equal, a setting with a high satisfaction score performs better than one with a low score.

Residential satisfaction can be considered an overall global evaluation of the residential environment from inhabitants' perspectives. Such a construct includes different aspects belonging to three main domains: the realm of people's affective response toward the residential environment, the realm of their cognitive evaluations of the same environment, and the realm of the action carried out in that environment. Very few studies have dealt with the behavioral component, whereas the first two components have been studied in greater depth. The first two aspects are grasped by two important constructs that are strictly related to, and partly overlapping with, residential satisfaction: place attachment and perceived residential environment quality (PREQ), respectively.

Place attachment, or more specifically residential attachment, refers to the feelings that people develop over time and come to experience with respect to their homes or their residential neighborhoods. It encompasses cognitive, motivational, and behavioral aspects, reflected in psychological correlates such as the tendency of people to give favorable evaluations of their dwelling places, reasons for improving them, and reluctance to leave them. Place attachment can be operatively defined in terms of inhabitants' answers to items such as (in the case of the residential neighborhood scale) "This neighborhood is part of me" and "It would be very hard for me to leave this neighborhood." The aim is to grasp people's overall affective responses to their own residential environments.

PREQ refers to people's evaluations of the degree of quality of a range of single features of their homes or neighborhoods. PREQ can be operatively defined in terms of inhabitants' answers to items such as (again referring to the residential neighborhood scale) "Buildings have unpleasant colors" and "The air is clean." The aim is to grasp specific cognitive evaluations with respect to a range of certain residential features from people's points of view.

Moreover, environmental quality is typically considered to be a multidimensional construct. Residential quality can include an array of features (e.g., from architectonic and town planning features to human and social features, from services to contextual elements). Each of these categories can include a range of subcategories, with each consisting of many different items.

Finally, environmental quality can be viewed in terms of both technical and observational assessments. In terms of the former, environmental quality is evaluated on the basis of technical systems of measurement (so-called objective or expert assessment). In terms of the latter, perceived environmental quality is evaluated on the basis of consensual impressions expressed in an ordinary language framework by the inhabitants (so-called subjective or layperson assessment).

2. URBAN ENVIRONMENT QUALITY IN EXPERTS' AND LAYPERSONS' EVALUATIONS

Thus, the quality of the very same urban environment can be evaluated from at least two different perspectives: the technical expert's assessment and the observer-based layperson's assessment. The former is often referred to as objective because it involves tools and measures (e.g., metrics, weights) of scientific and technical disciplines, that is, mechanical monitoring equipment or other physical means to produce a reading of environmental quality (it does not require individual exposition). The latter perspective may be referred to as subjective because it relies on self-report tools through which people express their own perceptions, observations, and impressions; that is, it employs the perceptual abilities of humans to judge environmental quality. This means that observational assessment offers a measure of the quality of the environment as it is experienced. However, both kinds of environmental assessments aspire to objectivity inasmuch as they strive to have reproducible measures that are valid, reliable, sensible, and useful. Similarly, they both share subjectivity because technical assessment relies on human decisions regarding which dimensions to examine and sampling times and places as well as on

interpreting results, whereas observational assessment by definition is derived from individual experience-based responses (although these responses sometimes can show strikingly high intersubjective agreement).

Both kinds of evaluation aim to evaluate the environment in which people live, but it is likely that different values, ideals, and goals concerning the environment underlie the expert's technical assessment and the layperson's observational assessment. Theoretically, such underlying discrepancies reflect the processes of social and cultural construction of the environment and of environmental issues, where social conflicts and controversies affect environmental descriptions and evaluations. For example, ethical, aesthetic, and ideological forces play a role in defining what should be considered as "polluted" or "dangerous" for the environment and for people.

Therefore, the degree of correspondence between the different types of evaluations (technical and observational) can be problematic. For example, when assessing the quality of urban green areas, experts and laypersons are often in strong disagreement because laypersons place strong importance on green area accessibility and usability, whereas experts' assessments tend to overlook these aspects and emphasize the quality of the biological element per se. Thus, different systems of values can generate different, and sometimes opposite, evaluations of the quality of the same urban residential environment. For this reason, it is important to compare and integrate technical assessments with observational ones in designing and managing the environment.

The following sections discuss the general argument of observational assessment by referring specifically to the residential neighborhood scale.

3. DESCRIBING RESIDENTIAL SATISFACTION AND PERCEIVED RESIDENTIAL ENVIRONMENT QUALITY

A basic issue in observational assessment is to establish criteria and devise valid and reliable tools to operatively describe and measure constructs such as residential satisfaction and perceived quality of the residential environment. The general aim here is to know how laypersons (i.e., residents) perceive and evaluate these aspects. Such a problem can generally be framed within a broader "place theory" that considers any place experience, at any given scale, as psychologically composed of three main elements: the physical attributes of the setting, the activities carried out in the setting, and the descriptions or conceptions that people have of the setting. Therefore, describing inhabitants' residential satisfaction and perceived residential quality is a prerequisite to obtaining a complete environmental–psychological picture of their residential place (which should, however, be complemented by a technical assessment of the place's physical features as well as by an assessment of inhabitants' uses of the place).

In describing how inhabitants perceive, conceive, and evaluate a residential place, two main methodological approaches can be adopted: a more "inductive" approach, which moves from how inhabitants spontaneously think and/or talk about their own residential environment, and a more "deductive" approach, which moves from the features of the residential environment that the researcher considers to be theoretically relevant and then elicits inhabitants' evaluations of these features. Often, the two approaches represent two successive phases of a broader research project.

On the basis of the first approach, inhabitants can take part in unstructured or semistructured interviews, or in small group discussions, and the researcher can then try to describe the way in which they talk about their residential environment. This approach usually privileges qualitative evidence to depict the commonsense ways in which to conceive and evaluate the residential environment.

On the basis of the second approach, inhabitants are typically administered questionnaires or checklists that present them with several items, referring to their residential environment, selected as relevant within the researcher's theoretical framework. Then, they are requested to respond to the items by filling in agreement–disagreement or satisfaction–dissatisfaction graduated scales. This approach typically privileges quantitative methods, gathering large samples of inhabitants and using statistics to infer the population's main trends.

The first approach can offer an in-depth and fresh description of the "residential ethnography," that is, inhabitants' ways of giving meaning to, and organizing, their social–physical residential reality. The second approach can offer quick, systematic, and broad information on how inhabitants position themselves on the whole array of features that the researcher considers to be crucial to evaluate the PREQ.

Each method can emphasize different aspects. For instance, consider perceived residential neighborhood

quality. If one concentrates on inhabitants' free expressions, one will probably notice that they often bring together the evaluation of aspects that are technically or theoretically considered separate (e.g., architectonic or urban planning features with social features, services with social features). On the other hand, if one asks inhabitants directly, through ad hoc prepared scales, to evaluate each feature of the residential environment, one will end up with separate pure indicators for each feature. In the first case, one can appreciate the fact that the method allows one to "reconstruct" the weft of inhabitants' residential perceptions and evaluations, discovering how they psychologically bring together different features on the basis that, for example, they all pertain to urban safety. In the second case, one can appreciate the possibility of knowing inhabitants' positions on exactly the same matters in which one is interested, and these can then be used to make systematic comparisons between different groups or places or between these perceptions and other expert-based assessments of the very same features.

Having said this, it must be recognized that the results of different studies carried out in several countries over the past 30 years or so converge to show a stable structure underlying people's evaluations of their residential environment. Turning again to the residential neighborhood scale, so far as neighborhood satisfaction is concerned, its description and measurement has traditionally been pursued in a very simple way: through a single item. On the other hand, so far as neighborhood environment quality is concerned, consistent evidence in the literature confirms that a theoretical model that includes three or four main areas can account for inhabitants' perceived neighborhood quality. Specifically, the first three areas are traditionally established in the literature, whereas the fourth area has been proposed in more recent contributions:

1. Spatial features (architectonic and town planning, including green areas)
2. Human and social features (population and type of social relations)
3. Functional features (available services)
4. Contextual features (lifestyle, pollution, maintenance, and care)

Each of these macro-areas consists of several aspects, and measurement scales are available to describe inhabitants' perceived quality of each specific aspect. The following is an example of a list of PREQ indicators for which specific valid and reliable scales exist (all referring to the neighborhood or the neighbors):

1. Building aesthetics
2. Building density
3. Building volume
4. Internal practicability
5. External connections (with the city)
6. Green areas (presence and care)
7. Discretion and civility
8. Security and tolerance
9. Sociability and cordiality
10. School services
11. Social care services
12. Sport services
13. Social–cultural activities
14. Commercial services
15. Transport services
16. Relaxing versus distressing pace of life
17. Stimulating versus boring pace of life
18. Environmental health (unpolluted environment)
19. Upkeep and care

4. EXPLAINING RESIDENTIAL SATISFACTION

In general, two main models have been proposed in the literature to explain residential satisfaction. A first model considers residential satisfaction as a criterion variable and seeks to identify its most relevant predictors so as to provide useful information for town planning and management. A second model uses residential satisfaction as a variable predicting other relevant urban behaviors such as residential mobility and residential choice.

Residential satisfaction can depend on different kinds of factors: personal, social, and physical. Among personal factors, residential satisfaction tends to increase with age, socioeconomic level, and improvement expectations. It tends to decrease when people are characterized by an aggressive personality trait or when people compare their residences with high-standard solutions.

Among social factors, residential satisfaction tends to increase with good neighbor relations, with similarity to neighbors, with a balance of separation and togetherness with neighbors that suits them, and with adherence to and conformity with social norms regarding "usual" housing arrangements for people like themselves.

Among physical factors, single-family dwellings are more satisfying than apartments (at least in North America), and green areas determine higher residential satisfaction. The preferred architectural style, as well as the interior, tends to vary with the cultural background of inhabitants. Color has complex relations with residential satisfaction as well, depending not only on hue but also on saturation and brightness. In general, residential satisfaction is greater when residences match those of inhabitants' own culture, probably because residences that reflect the culture are designed to complement behavior patterns typical of that culture.

Moreover, all of these factors can affect not only the level of residential satisfaction but also very specific preferences for residential solutions and choices at different scale levels.

In sum, within a general organizing framework or model, all of the previously mentioned factors can affect inhabitants' impressions of their residential places, and these in turn affect their cognitive, affective, and behavioral responses to their residences. These responses can also be affected (directly or indirectly) through the inhabitants' impressions by the objectively measurable features of their residences.

Again considering the example of inhabitants' cognition and affect at the residential neighborhood scale, it has been shown that neighborhood satisfaction and neighborhood attachment are related to a variety of neighborhood features, specifically to personal, social, and residential factors as well as to certain aspects of PREQ. So far as personal or residential features are concerned, age, home ownership and size, and length of residence in the neighborhood, for example, are usually positively related to neighborhood satisfaction and/or attachment, whereas home crowding is usually inversely related to it.

Among the features of PREQ, some are generally positively associated with residential satisfaction and attachment: general social and psychological climate (opportunities and quiet), building aesthetics and presence and maintenance of green areas, social relations (sociability, discretion, safety, and tolerance), and presence of, and access to, services and facilities. In the case of attachment, compared with residential satisfaction, the quality of social relations bears a stronger importance. However, both residential satisfaction and attachment are associated to a "cocktail" of quality indicators belonging to different areas: architecture and town planning features, social relations, services, and general context. That is, they both share a multicomponent or multidetermined nature.

5. APPLICATIONS

Residential satisfaction, residential attachment, and PREQ indicators can be used with applied aims in specific settings for residential design or evaluation. Typically, important applied research issues pertain to (a) the relationship between experts' and laypersons' environmental assessments, (b) the relationship between different evaluations (e.g., cognitive and affective, molecular and molar), and (c) the use of these constructs and measures in postoccupancy evaluation (POE) procedures.

The first two issues were mentioned in the opening paragraphs. So far as (residential) POE is concerned, it can be defined as an examination of the effectiveness for human users of occupied designed (residential) environments. It should be a systematic process, including all user groups, all important activities occurring in a place, and rigorous methods of collecting and analyzing data. It can refer to any relevant constructs and instruments, but when residential environments are being considered, it would certainly assess both residential satisfaction and some PREQ indicators and probably residential attachment as well. It would try to understand, for example, what is the degree of residential satisfaction and what are the environmental quality indicators—or other personal, social, residential, and physical factors—most associated with it for the inhabitants of the considered place.

In sum, knowing how residential environment quality (both perceived and technical) relates to residential satisfaction and residential attachment, with reference to the inhabitants of a specific place, can help to inform its design, management, and policies.

See Also the Following Articles

Environmental Design and Planning, Public Participation in ■ Environmental Stress ■ Extreme Environments and Mental Function ■ Person–Environment Fit ■ Restorative Environments ■ School Environments ■ Urban Environments and Human Behavior ■ Work Environments

Further Reading

Aragonés, J. I., Francescato, G., & Garling, T. (Eds.). (2001). *Residential Environments: Choice, Satisfaction, and Behavior*. Westport, CT: Bergin & Garvey.

Bonaiuto, M., Aiello, A., Perugini, M., Bonnes, M., & Ercolani, A. P. (1999). Multidimensional perception of residential environment quality and neighborhood

attachment in the urban environment. *Journal of Environmental Psychology*, *19*, 331–352.

Bonaiuto, M., Bonnes, M., & Continisio, M. (2004). Neighborhood evaluation within a multi-place perspective on urban activities. *Environment & Behavior*, *36*, 41–69.

Bonaiuto, M., Fornara, F., & Bonnes, M. (2003). Indexes of perceived residential environmental quality and neighborhood attachment in urban environments: A confirmation study on the city of Rome. *Landscape and Urban Planning*, *65*, 41–52.

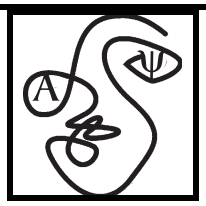
Bonnes, M., & Bonaiuto, M. (1995). Expert and layperson evaluation of urban environmental quality: The "natural" versus the "built" environment. In Y. Guerrier, N. Alexander, J. Chase, & M. O'Brien (Eds.), *Values and the environment: A social science perspective* (pp. 151–163). New York: John Wiley.

Craik, K. H., & Feimer, N. R. (1987). Environmental assessment. In D. Stokols, & I. Altman (Eds.), *Handbook of environmental psychology* (Vol. 2, pp. 891–918). New York: John Wiley.

Gifford, R. (2002). *Environmental psychology: Principles and practice* (3rd ed.). Colville, WA: Optimal Books.

Lawrence, R. J. (2002). Healthy residential environments. In R. B. Bechtel, & A. Churchman (Eds.), *Handbook of environmental psychology* (pp. 394–412). New York: John Wiley.

Tognoli, J. (1987). Residential environments. In D. Stokols, & I. Altman (Eds.), *Handbook of environmental psychology* (Vol.1, pp. 655–690). New York: John Wiley.



Restorative Environments

Terry Hartig
Uppsala University, Gävle, Sweden

1. Introduction
2. Essential Components of Theories about Restorative Environments
3. Extant Theories about Restorative Environments
4. Conclusions and Prospects
Further Reading

GLOSSARY

restoration The process of renewing, recovering, or reestablishing physical, psychological, and social resources or capabilities diminished in ongoing efforts to meet adaptive demands.

restorative environment An environment that promotes (and not merely permits) restoration.

stress A process of responding to an excess of demands relative to the resources needed to cope with those demands.

People inevitably deplete adaptive resources in everyday life. Persistent failure to restore needed resources will ultimately harm psychological and physical health. Restoration proceeds more effectively in some environments than in others due to a relative absence of demands as well as to qualities that promote restoration. Knowledge of the restoration-promoting qualities of environments can serve applied health fields.

1. INTRODUCTION

People inevitably deplete physical and psychological resources as they meet self- and externally imposed demands within changing environments. Failure to reestablish vital capabilities for effective action can harm health through multiple pathways. Some of those pathways have received close attention in research on the effects of chronic stress. As commonly defined, stress arises from an excess of demands relative to the resources needed to cope with those demands. This formulation implies that stress becomes chronic when excessive demands persist and the person can neither acquire the new resources needed to obviate those demands nor apply available resources more effectively. Less obviously, stress may also persist when the person cannot access an environment that supports sufficiently rapid or complete restoration of necessary resources diminished in the effort to cope.

In searching for the sources of chronic stress, researchers have identified a variety of social and physical environmental stressors such as crowding and noise. A relative absence of such demands may permit restoration. However, restorative environments warrant definition in positive terms rather than in negative terms. Environmental psychologists, in extending research on topics such as landscape aesthetics and psychological benefits of wilderness experience, have proposed qualities of environments that promote restoration. This theorizing has value for fields concerned with psychosocially

©2004 Elsevier Inc.
All rights reserved.

mediated relations between environment and health, such as health psychology, public health, social epidemiology, and psychosomatic and preventive medicine.

2. ESSENTIAL COMPONENTS OF THEORIES ABOUT RESTORATIVE ENVIRONMENTS

To better understand the current status and avenues for further development of theories about restorative environments, consider what such theories must include. First, a theory must specify some condition from which a person becomes restored. Second, it must describe the process of restoring some set of resources. Finally, it must characterize the environments that promote a process of restoration.

2.1. Antecedent Condition

Restoration cannot occur unless the possibility for restoration exists. Necessarily, before a person can become restored, he or she must have depleted some of those resources that are useful for maintaining and improving adaptation to the environment. Whether those resources are biological, psychological, or social, their availability is essential to continued adaptation. Over the long run, an inability to renew depleted resources may have grave consequences for effective action, subjective well-being, and physical health.

That the resources of interest here regularly and predictably become diminished helps one to distinguish the driving concerns of restorative environments theory from those of therapies and rehabilitative strategies. True, restorative environments applications can serve goals of therapy and rehabilitation defined by one health professional or another. Yet therapy often has to do with capacities that a person never had and, thus, that could become "restored" only with reference to normative criteria based on some population. Alternatively, therapy and rehabilitation may focus on capacities that a person has lost to an accident or some pathological process rather than in the normal course of adaptation to a continuously changing environment. This means that restorative environments theory has a broader reach than do therapy and rehabilitation in general. It does not necessarily exclude relatively unusual events that may reduce some capacity for action, but its area of concern extends to states that reflect normal "wear and tear."

2.2. Restorative Process

The term "restoration" denotes some set of processes through which one or more individuals renew or reestablish adaptive resources or capacities that have become diminished. Whether the resources of interest have a biological, psychological, or social character, the processes included under the restoration rubric here have in common a psychosocial character. Having specified an antecedent condition, a theory ought to define the given restorative process in terms of the resources that become restored and the psychological, physiological, and/or social mechanism(s) for their renewal. Furthermore, because all processes extend through time, a complete description of a restorative process will refer to temporal features and parameters such as stages in restoration and the time required for restoration. Measurements of appropriate variables at various time points provide evidence as to whether or not restoration is actually under way.

2.3. Environmental Context

All restoration occurs in the course of some activity, and all activity occurs in some environment. Restorative processes are not necessarily specific to a particular environment, but they may proceed more readily or smoothly in some activities and environments than in others. Those environments that promote (rather than merely permit) restoration can be referred to as "restorative." Whether with respect to their physical, social, activity, temporal, or spatial features, some form of contrast with another relatively demanding environment is implicit in the description of an environment as restorative.

2.3.1. Significance of the Natural–Urban Distinction

One set of contrasts aligns with a coarse distinction between natural and urban environments. In studying restorative environments, some researchers have focused on people living in cities who occasionally go to relatively natural environments for restoration. Convergent practical concerns motivate this focus. First, most people today reside in urban areas. Conditions common in urban areas can impose heavy demands on people. Second, it might not be easy to ensure that people who live in cities have easy access to natural environments. Other people may prefer to close off access to those environments and disturb their natural character in the pursuit of economic gain.

Although these problems are related, work on them within environmental psychology has, for the most part, proceeded independently. A large body of research attests to harmful effects of noise, crowding, air pollution, and other demands in urban areas. Another large body of research describes psychological and social amenity values of natural areas placed at risk by urbanization and the extraction of timber, minerals, and so forth. In restorative environments research, the practical and theoretical connections between these bodies of work become distinct.

3. EXTANT THEORIES ABOUT RESTORATIVE ENVIRONMENTS

Two theories have guided most of the research on restorative environments to date. Both are rooted mainly in research on the psychological values of natural environments, but in specifying an antecedent condition from which people might need restoration, both have referred to work performed under the general "environmental stress" rubric.

3.1. Directed Attention Restoration

Building on some 30 years of research on environmental cognition, environmental preferences, and the psychological benefits of nature experiences, Stephen and Rachel Kaplan developed a theory concerned with the capacity for directing attention.

3.1.1. Antecedent Condition

The Kaplans assumed that a person's ability to direct attention depends on a central inhibitory capacity. To focus on something that is not of itself interesting, the person must inhibit competing stimuli that are more interesting. Doing so requires effort, and with prolonged or intensive use, the person's ability to inhibit competing stimuli will diminish or become fatigued. The Kaplans described a variety of negative consequences that the person may suffer when this inhibitory capacity diminishes, including irritability, failure to recognize interpersonal cues, reduced self-control, and increased error in performance of tasks requiring directed attention. Their account of directed attention fatigue has much in common with accounts of cognitive effects and aftereffects of stress or informational overload.

3.1.2. Restorative Process

According to attention restoration theory, a person can restore a diminished capacity for voluntarily directing attention when he or she experiences fascination, a mode of attention that the Kaplans assumed to have an involuntary quality, to be effortless, and to not have capacity limitations. When a person can rely on fascination in ongoing activity, demands on the central inhibitory capacity are relaxed and a capacity for directing attention can be renewed.

Regarding the time course of attentional restoration, the Kaplans proposed that a restorative process may continue through several stages, ranging from clearing one's head of random thoughts and cognitive "clutter," to renewing directed attention capacity, and ultimately to reflecting on matters of personal importance.

3.1.3. Environmental Context

As described by the Kaplans, fascination is engaged by objects or events or by processes of exploring and making sense of an environment. Yet fascination is not sufficient for restoration. Attention restoration theory also specifies being away, or getting psychological distance, from the work one usually does and from the pursuit of given goals and purposes—hence, from further demands on directed attention. Another factor is extent, that is, the sense that a physical or conceptual environment available for restoration is sufficiently large in scope to entertain continued exploration and is sufficiently coherent so that one can make sense of what is seen going on around himself or herself and relate it to some larger frame of reference. A fourth factor, compatibility, encompasses a person's inclinations, environmental supports for his or her activities, and environmental demands. It rests on the match among what a person wants to do, what the person can do, and what the person must do in the given environment. The Kaplans proposed that high compatibility allows for deeper levels of restoration.

Although many environments might afford the experience of being away, fascination, extent, and compatibility, the Kaplans argued that natural environments should do so more readily than do other environments. For example, natural environments may afford being away more readily due to a scarcity of reminders about work demands and a relative absence of people (with whom interactions may require directed attention). The Kaplans also asserted that natural environments are rich in aesthetically pleasing features, such as

scenery and sunsets, which evoke moderate or "soft" fascination that permits a more reflective mode.

3.1.4. Empirical Tests

Several quasi- and true experiments have tested the proposition that experiences of natural environments should promote better restoration of a capacity to direct attention than should experiences of other environments. In these studies, the researchers have operationalized directed attention capacity in terms of performance on tasks that require a participant to focus attention. For example, Hartig and colleagues reported a field experiment in which differential proofreading performance was seen after 40 minutes spent in either a nature reserve, a city center, or a passive relaxation condition. On average, posttest proofreading performance of participants randomly assigned to the natural environment condition was better than that of participants assigned to the other two groups.

3.2. Psychophysiological Stress Recovery

A second theory, developed by Roger Ulrich, focuses on patterns of affective and aesthetic response to visual stimulus characteristics of an environment. Of the range of possible reactions to what one sees in an environment at a given moment, Ulrich had a particular interest in those that promote psychophysiological stress reduction.

3.2.1. Antecedent Condition

Ulrich defined stress as a process of responding to a situation perceived as demanding or threatening to well-being. He further assumed the operation of an evolved system for directing behavior in situations relevant to continued survival. That system depends on "hard-wired" affective responding in the selection of a behavioral strategy (i.e., approach or avoidance) and the simultaneous mobilization of physiological resources needed to execute that strategy. Stress becomes manifest in self-reports of negative emotions and short-term changes in physiological systems that indicate negative affect and heightened autonomic arousal.

3.2.2. Restorative Process

This theory focuses primarily on restoration from stress as a potential mode of affective responding to a visual stimulus array. It proposes that restoration can occur when a scene elicits feelings of mild to moderate interest, pleasantness, and calm. For a person who enters a situation experiencing psychophysiological stress and needing to renew resources for further activity, it could be adaptive to continue viewing the scene in a nonvigilant manner. While the scene is being viewed, negative affects are replaced by positive ones, the person's interest is held, and physiological arousal declines.

3.2.3. Environmental Context

It is believed that such restorative responses are initiated very rapidly by the perception of certain visual patterns in the environment. Ulrich described these as aspects of the visual stimulus array that might not provide enough information for conscious cognitive judgments but that, nonetheless, can effectively elicit a generalized affective response. He proposed that qualities of the visual stimulus array, such as moderate depth, moderate complexity, and the presence of a focal point, will promote restoration. Nature enters into this theory as well in that particular environmental contents—vegetation and water—may rapidly evoke positive affective responses. People may be biologically prepared to quickly acquire and retain a liking for environmental features that would have been significant for the survival of our early ancestors. These include water, uniform grassy ground cover, and the forms and distribution of trees characteristic of savannah landscapes—a setting of human evolution.

3.2.4. Empirical Tests

Experiments guided by this theory have involved 10- to 18-minute photographic simulations of natural and urban environments and have documented differential change in emotional and physiological outcomes measured during or immediately after the period of the simulation. For example, Ulrich and colleagues reported a study in which participants viewed a stressful industrial accident film followed by a 10-minute video of either a natural setting, urban traffic, or an outdoor pedestrian mall. Poststressor recovery trajectories for frontalis muscle tension, skin conductance, heart period, and pulse transit time differed as a function of the environment viewed, with nature simulations promoting the fastest returns toward baseline and the lowest overall levels. Changes in self-reported affect converged with the physiological results in showing greater restorativeness with the nature videos.

4. CONCLUSIONS AND PROSPECTS

The theories just reviewed propose a variety of restorative qualities of environments, including complexity, depth, and focality in the visual stimulus array as well as being away, fascination, extent, and compatibility in the experience of the purposive actor. Both theories emphasize the relative restorative power of some natural environments, assuming that people remain adapted, to some degree, to the environments of our prehistoric ancestors and observing that those evolutionary settings would contrast sharply with today's demanding urban environments. Research guided by these theories has found, with considerable consistency, relatively greater restorativeness of commonplace, nonthreatening natural environments as compared with urban outdoor environments.

4.1. Prospects for Theory and Empirical Research

One perspective on the future of restorative environments research encompasses the possibilities for working with the two theories just overviewed. First, their claims about the mechanisms at work in restoration may well need refinement or revision in the light of independent developments in other areas of psychological inquiry. Second, tests of each theory can more closely consider attributions about the restorative qualities they propose. For example, to what extent does fascination versus being away account for a possible difference in attentional effects of sampled urban and natural environments? Third, studies can test predictions about how a given restorative process leads to some novel outcome, as was done in Kuo and Sullivan's study of directed attention fatigue and aggression by residents of urban low-income housing with barren versus green surroundings. Fourth, the moderating role of individual differences in attention restoration and stress recovery remains little studied by environmental psychologists. Just as individuals consistently distinguish themselves in the face of demands, as with the heightened physiological reactivity of a repressive coping style, so too may individuals show distinctive patterns of restoration. To take a final example, future work can further explore the complementarity of the two theories. The theories appear to deal with processes that can run simultaneously, albeit not for the same duration. They specify antecedent conditions that might not necessarily be related to one another; that is, a reduced capacity for directing attention may occur independently of the elevated physiological arousal characteristic of stress and vice versa. And although both of the theories note how particular environmental configurations hold attention, they then diverge in following the implications of this for different resources, namely, directed attention capacity versus physiological response capabilities.

The prospects for restorative environments research involve more than further work with the extant theories. The general theoretical framework provided at the outset opens a broader perspective. It implies possibilities for theoretical extension through reference to other antecedent conditions, restorative processes, and environmental variables. For example, the theories reviewed previously have rather little to say about how social factors might work in promoting (vs permitting) restoration. Noting this gap, Staats and Hartig discussed how the presence of another person can enable and enhance restoration in a given environment. Another person can enable restoration by making it possible for an individual to enter an environment of high restorative quality, for example, by providing protection. Issues of access aside, once in a given environment, another person might enhance the restorative quality of the environment by providing the individual with welcome observations and interpretations that increase interest in the environment.

To take another example, consider the cumulative effect of repeated restorative experiences on a specific person's relationship to a given place. Recognizing that people do commonly establish such relationships, Korpela and colleagues proposed that a person becomes attached to places relied on for restorative experiences and that such attachments, in turn, contribute to that person's identity. Furthermore, a place may become more restorative for that person due to the personal bonds. Thus, restorative experience, place attachment, and place identity become reciprocally influential.

4.2. Practical Prospects

Beliefs about restorative qualities of environments have appeared over millennia and across cultures. Some such beliefs, expressed in paradisiacal religious imagery, arguments for wilderness preservation, claims about the values of national and urban parks, and philosophies of institutional health care, have attributed restorative quality to natural scenery and features such as vegetation and water.

Recent practical efforts have sought to capitalize on the growing body of empirical research on the restorative values of nonthreatening natural settings and scenery in combination with improved understanding of the biological and behavioral pathways through which chronic stress harms health. At the scale of buildings and their immediate surroundings, architects, interior designers, and landscape architects have used "healing gardens," landscape paintings and photographs, and window views of natural elements to make highly technological, (often) anxiety-arousing health care settings more humane. Design measures sensitive to user needs for restoration may translate into improved patient outcomes such as less pain and shorter stays. They may also yield benefits to worried family members and hard-working staff.

At the scale of neighborhoods and cities, local decision makers weighing the value of competing land uses can consider increasingly well-grounded claims that green areas serve health. A well-articulated urban green structure can provide urban residents with easily accessible opportunities for exercising and for interrupting the stress process during the course of everyday life in the city. That said, visual access to natural scenery through windows in residences and workplaces may provide restorative benefits independent of physical activity. The increment paid for a pleasant view from a dwelling or hotel room speaks to the value that people attach to such benefits.

People value views with other than natural content, to be sure, and not all restorative environments are natural environments. By considering the potential restorative value of human environments in general, and translating the knowledge secured into policy, planning, and design measures, one can better support adaptation and, thus, promote health and well-being.

Acknowledgments

The author thanks Florian Kaiser, Kalevi Korpela, and Henk Staats for comments on an earlier version of this article.

See Also the Following Articles

Conservation Behavior ■ Environmental Assessment ■ Environmental Design and Planning, Public Participation in ■ Environmental Psychology, Overview ■ Pro-environmental Attitudes and Behavioral Change ■ Residential Satisfaction and Perceived Urban Quality ■ Urban Environments and Human Behavior

Further Reading

Cimprich, B., & Ronis, D. L. (2003). An environmental intervention to restore attention in women with newly diagnosed breast cancer. *Cancer Nursing*, *26*, 284–292.

Hartig, T. (2001). Guest editor's introduction [to special issue on restorative environments]. *Environment and Behavior*, *33*, 475–479.

Hartig, T., Evans, G. W., Jamner, L. D., Davis, D. S., & Gärling, T. (2003). Tracking restoration in natural and urban field settings. *Journal of Environmental Psychology*, *23*, 109–123.

Hartig, T., Mang, M., & Evans, G. W. (1991). Restorative effects of natural environment experiences. *Environment and Behavior*, *23*, 3–26.

Hartig, T., & Staats, H. (2003). Guest editors' introduction: Restorative environments. *Journal of Environmental Psychology*, *23*, 103–107.

Hartig, T., Johansson, G., & Kylin, C. (2003). Residence in the social ecology of stress and restoration. *Journal of Social Issues*, *59*, 611–636.

Herzog, T. R., Black, A. M., Fountaine, K. A., & Knotts, D. J. (1997). Reflection and attentional recovery as two distinctive benefits of restorative environments. *Journal of Environmental Psychology*, *17*, 165–170.

Kaplan, R. (2001). The nature of the view from home: Psychological benefits. *Environment and Behavior*, *33*, 507–542.

Kaplan, R., & Kaplan, S. (1989). *The experience of nature: A psychological perspective*. New York: Cambridge University Press.

Kaplan, S. (1995). The restorative benefits of nature: Toward an integrative framework. *Journal of Environmental Psychology*, *15*, 169–182.

Kaplan, S., & Talbot, J. F. (1983). Psychological benefits of a wilderness experience. In I. Altman, & J. F. Wohlwill (Eds.), *Human behavior and environment: Advances in theory and research*, Vol. 6: *Behavior and the natural environment* (pp. 163–203). New York: Plenum.

Knopf, R. C. (1987). Human behavior, cognition, and affect in the natural environment. In D. Stokols, & I. Altman (Eds.), *Handbook of environmental psychology* (Vol. 1. pp. 783–825). New York: John Wiley.

Korpela, K. M., Hartig, T., Kaiser, F. G., & Fuhrer, U. (2001). Restorative experience and self-regulation in favorite places. *Environment and Behavior*, *33*, 572–589.

Kuo, F. E., & Sullivan, W. C. (2001). Aggression and violence in the inner city: Impacts of environment via mental fatigue. *Environment and Behavior*, *33*, 543–571.

Staats, H., & Hartig, T. (2004). Alone or with a friend: A social context for psychological restoration and environmental preferences. *Journal of Environmental Psychology*, *24*, 199–211.

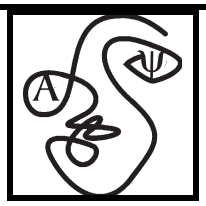
Ulrich, R. S. (1983). Aesthetic and affective response to natural environment. In I. Altman, & J. F. Wohlwill (Eds.), *Human behavior and environment: Advances in theory and*

research, Vol. 6: *Behavior and the natural environment* (pp. 85–125). New York: Plenum.

Ulrich, R. S. (1984). View through a window may influence recovery from surgery. *Science*, *224*, 420–421.

Ulrich, R. S. (1993). Biophilia, biophobia, and natural landscapes. In S. Kellert, & E. O. Wilson (Eds.), *The biophilia hypothesis* (pp. 73–137). Washington, DC: Island Press.

Ulrich, R. S., Simons, R., Losito, B. D., Fiorito, E., Miles, M. A., & Zelson, M. (1991). Stress recovery during exposure to natural and urban environments. *Journal of Environmental Psychology*, *11*, 201–230.



Risk-Taking and Sexual Behavior

Vita C. Rabinowitz
Hunter College, New York, New York, USA

1. Defining Sexual Risk-Taking
2. What Are Risky Sexual Behaviors?
3. What Groups Are at High Risk for Teen Pregnancy and STIs?
4. What Social, Interpersonal, and Personality Factors Are Associated with Risky Sexual Behaviors?
5. Why Do People Engage in Risky Sex?
6. How Can We Reduce the Incidence of Risky Sex?
Further Reading

GLOSSARY

acquired immunodeficiency virus (AIDS) A condition caused by the human immunodeficiency virus and characterized by destruction of the immune system, which leaves the infected person vulnerable to a host of serious opportunistic diseases.

human immunodeficiency virus (HIV) The virus that causes AIDS.

human papilloma virus The virus that causes warts on the anus or genitals.

risky or unsafe sexual behaviors Behaviors that make it more likely that one will contract a sexually transmitted infection or that will contribute to an unwanted pregnancy, especially for adolescents.

sexually transmitted infection (STI) Any of a number of parasitic, bacterial, or viral infections that can spread through sexual contact.

How do you choose your sexual partner(s)? When do you initiate, accept, and reject sexual advances? When should you and your partner become sexually intimate, and what kinds of sexual practices will you engage in? What factors into your decisions about condom use? When should you get tested for HIV infection, and when should your partner be tested? What sexual risks do you take, and when and why do you take them?

Like eating and sleeping, sex is a natural function, but alone among the natural functions, sex is widely regarded as risky to people's moral, physical, and psychological well-being. Of the natural functions, it is the one most often feared, proscribed, regulated, and punished. Indeed, many of us learn that sex is dangerous before we learn that sex is natural, pleasurable, or beautiful. This article discusses what psychologists have learned about sexual risk-taking. First, sexual risk-taking and its possible negative consequences are defined. Then, the factors that make sex more or less risky are identified. This is followed by a discussion of some groups who are at special risk for negative sexual outcomes and the contexts or conditions that increase sexual risk-taking. Finally, a discussion of what psychological research can teach us about how to make more responsible sexual decisions is presented.

1. DEFINING SEXUAL RISK-TAKING

Although sex is a natural function, motives for sexuality, expressions of sexuality, and the consequences of sexual behavior vary widely within and among individuals, societies, and cultures, and change over time. All societies have an incest taboo, but masturbation,

©2004 Elsevier Inc.
All rights reserved.

prostitution, premarital sex, and polygamy, among other sexual activities, are condemned in some societies, tolerated in others, and championed in still others. In 1989, evolutionary psychologist Buss conducted a major survey of more than 10,000 participants from 37 cultures on six continents on the characteristics that people sought in mates. He found that the value that people place on virginity in prospective mates ranged from "indispensable" in China to "irrelevant" in Sweden and the Netherlands. Within societies, the meaning of a sexual act often depends on the context in which it takes place. For example, when access to women is severely restricted, as in male prisons, men having sex with men is not considered deviant and certainly does not cause these men to identify themselves as gay. Just as the meaning of a sexual act depends on many factors, what constitutes a sexual risk also depends on many factors and varies with societal and interpersonal contexts. After the sexual revolution of the 1960s and 1970s, the threat of AIDS in the 1980s has brought with it a return to more conservative, less risky forms of sexual behavior among many groups.

Sex can be risky because it threatens people's social status, emotional or psychological well-being, or physical health. In some societies, sex outside of marriage can threaten a woman's social standing as a wife and mother (if not also her life). Sexually intimate relationships are often consuming and complex, and they offer extraordinarily rich opportunities for feelings of disappointment, frustration, heartbreak, and loss. However, the main focus of psychological research on sexual risk-taking has been on unintended pregnancies and sexually transmitted infections, or STIs. For most psychologists who study human sexuality, risky sex refers to those sexual behaviors that render people vulnerable to threats to their physical health or well-being or that render women and girls, especially teenage girls, vulnerable to unintended pregnancies.

Psychologists' current emphasis on sexually transmitted infections and teen pregnancies in part reflects societies' heightened concerns about these conditions, especially since the latter part of the 20th century. Nothing has provided more impetus for the study of risky sex than the epidemic of AIDS, which is associated with sexual activity and threatens the lives of significant portions of humanity. Because it is sexually transmitted, incurable, lethal, and rapidly spreading, AIDS has garnered more attention worldwide than any new disease in recent history.

Teen pregnancy has long been regarded as a problem for society because teens, especially unmarried teens, are seen as incapable of providing children with the care and resources that they need in today's world. Although the rates of pregnancy, births, and abortions among girls aged 15–19 began to decline in the industrialized world, including the United States, in the early to mid-1990s, teen pregnancy continues to be a very real public health problem. Its rate in the United States remains much higher than that in most other developed countries. The consequences of unplanned teen pregnancies, and 9 in 10 pregnancies among unmarried teens are unplanned, can devastate young mothers, fathers, their children, and society at large. Unmarried teenage mothers are more likely to quit school, live in poverty, and go on public assistance than their peers. As Fraser *et al.*'s 1995 research suggests, they are more likely to have medical complications during pregnancy and childbearing than any group of fertile women except those in their late forties. They suffer high rates of miscarriage, stillbirths, birth complications, and infant mortality. The children of teenage mothers are at greater risk of physical, emotional, behavioral, and intellectual problems. They are more likely than other children to experience poor nutrition and health care; to suffer from family instability, abuse, and neglect; and to perform poorly in school.

Aside from the problems of teen pregnancy, childbearing, and parenting per se, adolescent sexuality raises other concerns. Studies show that approximately 30% of adolescent women do not use any form of contraception during their first sexual intercourse, and fewer than half of sexually active girls use condoms every time they have vaginal intercourse. Because sexually active teenagers use condoms inconsistently, if at all, teens are exposed to a wide range of STIs, including AIDS. According to the Centers for Disease Control and Prevention (CDC), in 2000, young people were among the fastest growing groups of those infected with HIV. Twenty percent of people in the United States who are diagnosed with AIDS are in their twenties, and most of them were infected in their teens. Among teens and young adults, African American youth account for two-thirds of all cases of HIV/AIDS in people younger than age 25. Research reviewed by Rathus *et al.* suggests that adolescent girls are particularly vulnerable to STIs such as HIV/AIDS because the cervix at their age is not fully mature and is thus more susceptible to infection.

Sexual behavior among adolescents has increased more or less steadily throughout the past century. Today, approximately 60% of high school seniors, both girls and boys, have had sexual intercourse at least once, and girls and boys are having intercourse at increasingly younger ages. Indeed, many scientists and laypeople regard all sex among young teens as inherently risky, no matter what forms of contraception and STI prevention are used. Many young people are unprepared for the complex consequences and responsibilities that come with sexual relationships and suffer negative consequences for sexual expression. As Rathus *et al*. note, there is evidence that girls who have sex early, particularly those who become pregnant and have children, exhibit lower self-esteem and are more likely to be depressed than their peers. Girls are often pressured into sexual activity by boys or their female peers, yet, at the same time, condemned for being sexual, especially outside of a committed relationship. This makes it difficult to understand how sexual expression for many girls is a freely chosen, empowered act. In fact, girls who become sexually active early often do not do so voluntarily. Younger teenage girls, especially those younger than age 14, are highly likely to be coerced into sexual intercourse by older males. Tragically, the sexual use of young girls by older men is common in some developing countries with high rates of HIV/AIDS, where sex with very young girls, particularly virgins, is seen as a way to avoid or even "cure" AIDS.

Despite the understandable concern about sexual decision making and sexual coercion among young people, some experts on adolescent sexuality oppose the presumption that adolescent sexuality is invariably problematic. Both Fine and Tolman, for example, view sex among teens as developmentally normal and appropriate rather than inherently pathological. They decry the extent to which schools, parents, and religious institutions unwittingly conspire to scare young people, especially young girls, away from sex by equating sexuality entirely with negative outcomes—violence, victimization, pregnancy, disease, immortality, and low-class behavior. In their view, sexual desire and exploration are vital parts of the developmental processes of adolescence for girls and boys. They cite the high rates of adolescent sexual behavior throughout the world, and the increasingly responsible sexual behavior among key groups of teens in the United States, as evidence for this. For most people, a sexual life begins in adolescence, and there is increasing evidence in the United States and the rest of the industrialized world that teens are not getting into as much trouble as they did in the recent past because of their sexual activity.

2. WHAT ARE RISKY SEXUAL BEHAVIORS?

As noted earlier, sex with another person carries with it a variety of potential risks. However, for the purpose of this analysis, risky or unsafe sexual behaviors are those that make it more likely that one will contract a STI or will contribute to an unwanted pregnancy, especially for adolescents. These include, but are not limited to, (a) having vaginal and especially anal sexual intercourse without a latex condom; (b) having multiple sex partners; (c) having sex outside of a committed relationship, especially with strangers, casual acquaintances, or people who have used intravenous drugs; (d) having sex under the influence of alcohol or other drugs, when sexual decision making is impaired; and (e) having sex under coercive conditions, when the possibility of meaningful choice, including the option of practicing safe sex, is eliminated.

3. WHAT GROUPS ARE AT HIGH RISK FOR TEEN PREGNANCY AND STIs?

It needs to be stated at the outset that it is sexual behavior that is risky, not being a member of a group. Being young or gay per se does not put a person at risk for STIs, for example, but engaging in anal sex without a latex condom does. That said, researchers have uncovered several correlates of risky sexual behavior, some of them involving subgroups of people. The fact that some geographic, social, ethnic, or other characteristics of people are associated with higher rates of teen pregnancies or STIs means only that some variables are statistically related and does not mean that there is anything inherent in group membership that causes these behaviors or negative outcomes.

Young people are vulnerable. Sexually active teens, especially young, unmarried teens, are generally at heightened risk for both unplanned pregnancies and STIs. This is in part because teens have less information about consequences of sexual behavior, less access to contraceptives, less skill in sexual communication and negotiation, and less ability to control a sexual

situation. As Rathus *et al.* suggest, another factor—the sense that young people have that they are invulnerable and can live forever—may also be at work. Many adolescents have the attitude that bad things, including pregnancies, STIs, and even death, simply cannot happen to them. However, as the Kaiser Foundation reported in 2002, more than 60% of new HIV infections in the United States occur in people aged 13–24.

People who have sexual practices involving sharing blood, semen, or feces are at high risk for HIV and other health problems. In the United States and much of the industrialized world, the AIDS epidemic first manifested in the 1980s among gay men and intravenous drug users. According to the CDC, in the United States AIDS predominantly affects men who have sex with men (50%) and people who share needles when they inject drugs (25–30%). These groups are at risk because some of their members engage in practices that involve sharing blood, semen, or feces, all of which can contain HIV. Researchers have found that gay men are more likely than lesbians and straight men and women to practice sexual behaviors that efficiently transmit the virus, for example, those in which semen containing the virus is directly transferred into the bloodstream through tears and open sores in the anus or rectum. These include practices that cause fissures, such as anal–genital sex and fist–anal activity (fisting). HIV is also contained in feces, so analingus or oral–anal contact (rimming) is a high-risk activity. Beyond specific sexual practices, gay men are also more likely than others to have multiple sex partners.

Many changes in the gay community have occurred since the onset of AIDS, and many gay men have changed their behavior to prevent contracting or spreading AIDS. However, although male–male anal intercourse has been on the decline since the epidemic took hold, there has recently been an increase in one particularly dangerous form of anal intercourse, "barebacking," which is anal intercourse without a condom. Unprotected receptive anal intercourse is especially risky because if the insertive partner is infected, the virus contained in the semen easily enters the bloodstream through the many small tears in the colon caused by anal penetration and thrusting. Anal sex is also associated with hepatitis B infection, anorectal human papilloma virus, and cancer. Again, it is behavior and not group membership that puts people at risk: Many heterosexual women are the receptive partners in anal sex, and, if their partners do not use latex condoms, they are at high risk of HIV infection.

Constellations of geographic, cultural, religious, and economic conditions have formed to make sex much more risky in certain places and times than others. HIV/AIDS is a global epidemic, but sub-Saharan Africa, which has approximately 10% of the world's population, bears the burden of approximately 66% of the world's AIDS cases. As Kelly notes, in some African countries one in four adults is infected, and it is estimated that 35% of boys now aged 15 will ultimately die of AIDS. However, in Africa, AIDS is mainly spread through unprotected vaginal intercourse and perinatally (i.e., around the time of birth) from mother to infant. Heterosexual transmission in Africa is facilitated by high rates of female genital mutilation among women, cultural practices and beliefs that discourage condom use, dry sex—sex without lubrication, migratory work patterns in men, and possibly low rates of circumcision among men. In Africa and, increasingly, in southern and Southeast Asia, men, women, and children are affected by AIDS in staggering numbers. The rates of both infection and death from infection are much higher in the poorer nations of the world. In the developed world, AIDS is still largely a disease of gay men and intravenous drug users, although heterosexual women infected by their male partners are the fastest growing group.

Within wealthy nations such as the United States, urban centers are the epicenter of the epidemic, with poverty, unemployment, alcohol use, and intravenous drug use all contributing to the problem. In the United States, race is confounded or correlated with many other social and economic factors. The Kaiser Foundation found that although African Americans make up only 13% of the U.S. population, they account for approximately 60% of new HIV infections and 54% of deaths from AIDS.

Women are at higher risk for some STIs, including HIV/AIDS, than men. There are many ways, from the biological to the social, in which sexuality is made more problematic for females than males. Of course, among the biological factors, one stands out above all others: Sexual intercourse is far more consequential for females than males because females alone bear children. However, being female also renders one more vulnerable to a variety of STIs. According to recent estimates, males transmit the HIV virus much more efficiently—approximately 10–18 times more efficiently—to females than females transmit the virus to males. Thus, acts of unprotected vaginal intercourse are in fact much more hazardous to women's health than they are to men's health.

4. WHAT SOCIAL, INTERPERSONAL, AND PERSONALITY FACTORS ARE ASSOCIATED WITH RISKY SEXUAL BEHAVIORS?

There are a number of factors that can predispose some people toward high-risk behaviors. One of the most reliable of all predictors of unsafe sex is the use of alcohol and other noninjectable drugs before engaging in sexual relations. Many studies show that substance use disinhibits people sexually and facilitates having multiple sex partners, sex with strangers, and sex without condoms. Teens who binge drink—who have five or more drinks on one occasion—are three times less likely to use a condom during sex than those who do not.

Relationship context matters. Women who believe that they are in monogamous relationships tend to view those relationships as safe, and they do not insist on condoms during sexual intercourse, regardless of what or how much they know about their partners' sexual histories or, for that matter, their partners' current sex lives outside of the primary relationship. Getting women in such relationships to see their behavior as risky means that they need to consider the possibility that their partner may not be "safe."

There is evidence that perceptions of the physical attractiveness of one's potential sex partner affect willingness to have unsafe sex. The more attractive a potential partner is perceived to be, the more likely people are to state an intention to have unprotected sex with that partner. This finding is robust and has been found among gay and straight respondents. It suggests that other desirable features of a partner besides physical attractiveness may lead people to abandon safe sex practices, for example, wealth or high social status.

Some personality factors appear to be related to sexual risk-taking. Self-regulation, defined as the ability to regulate one's emotions, attention, and behavior, is positively correlated with safe sex practices. On the other hand, people who are high in thrill or sensation-seeking are less likely to practice safe sex. For example, within the gay community, a minority of men who have sex with men deliberately seek out known high-risk behaviors, such as barebacking.

Despite all of the information available about safe sex, many people whose self-esteem is low do not accurately perceive risky behavior as such and consistently underestimate their vulnerability to diseases and conditions. Acknowledging self-destructive behavior is anxiety provoking, is damaging to one's view of oneself, threatens self-esteem, and possibly is threatening to a valued sexually intimate relationship as well. To avoid negative feelings about the self while continuing to enjoy the rewards associated with sexual behavior, some people deny that they are engaging in high-risk behavior; indeed, some people deny that they are sexually active at all. These people use an avoidant coping style that enables them to continue engaging in the risky behavior. Some studies show that women who have been victims of emotional, physical, and sexual abuse and have low self-esteem are particularly vulnerable to having risky sex.

Finally, some people have personal beliefs and attitudes about sexual behavior that make it difficult to practice safe sex. Some of them are deeply rooted in religious teachings or cultural practices. For example, the Catholic Church does not condone the use of any kind of birth control, including condoms, and many observant Catholics follow this teaching. In some cultures, particularly ones in which gender roles are rigidly enforced, women have little choice in their sexual expression and must follow the lead of their men, who may regard the use of a condom as an affront to their masculinity, a sign that their partners have been unfaithful, are sick, or do not love or trust them, and so on. Some men regard sex with condoms as less spontaneous or sexually arousing and, regardless of their or their partners' HIV status, insist on "skin-to-skin contact."

5. WHY DO PEOPLE ENGAGE IN RISKY SEX?

As already noted, sex is a natural function, and one that perpetuates the species. Sex is also one of the most rewarding experiences available to people. Given these fundamental truths, we would not appear to need further explanation for why people have sex, even risky sex. However, recent psychological research suggests that sexual behavior is not simply motivated, and that sexual behaviors, whether risky or safe, serve a wide range of psychological functions that have little to do with perpetuating the species, health protection, or disease avoidance. As Cooper *et al*. argue, people have many complex motives for having sex, and these motives need to be understood before effective interventions to decrease risky behaviors can be implemented. Specifically, Cooper *et al*. argue that having sex to meet different needs may be associated with distinctive patterns of

risky behavior. These investigators have developed a motivational model of sexual behavior that considers two dimensions: whether the motivation is the pursuit of positive experiences versus the avoidance of negative experiences, and whether the focus is the self or the social environment. The crossing of these two dimensions yields four constellations: using sex to escape or avoid threats to self-esteem (coping motives), using sex to avoid or escape negative social experiences (approval motives), using sex to enhance one's own positive emotions or experiences (enhancement motives), and using sex to enhance social connections with others (intimacy motives). Research suggests that these motives do indeed account for sexual behavior and predict sexual risk-taking, particularly when relationship variables and sexual self-esteem are taken into account.

6. HOW CAN WE REDUCE THE INCIDENCE OF RISKY SEX?

Sexual behavior is extremely complex, satisfies many motives, and is affected by many forces. Effective strategies for encouraging safe sex among young people not yet sexually active, and inducing people who practice risky sex to change their behavior, will need to include both individual and societal efforts. As McAnulty and Burnette argue, ultimate success will require a range of efforts involving the health care system, the government, the schools, religious organizations, peers, parents, lovers, and friends. Everything we know about changing personal behavior from smoking to exercising and eating suggests that changing sexual behavior will not be easy. Like these other personal behaviors, many risky sexual behaviors have the seductive feature of being rewarding in the short term but damaging in the long term. If anything, sexual behaviors are possibly more resistant to change than these other behaviors precisely because they are more complex, inherently interpersonal, highly rewarding, and private. McAnulty and Burnette suggested a number of behavioral patterns that can reduce or eliminate risky sexual practices.

6.1. Practicing Abstinence and Monogamy

The only prevention method that is 100% effective in eliminating the threats of unwanted pregnancy or STIs is abstinence. Abstinence can mean many things, from total avoidance of sexual activity to avoidance of penetrative sex. Many sexual activities that involve others, including kissing, petting, mutual masturbation, and the use of sex toys, are considered completely safe.

The next most effective STI prevention method is monogamy—having a sexually exclusive relationship with one partner. Obviously, for monogamy to be effective in preventing STIs, both partners must be free of disease. To prevent pregnancy, some form of birth control must be used.

6.2. Making a Commitment to Practice Safe Sex

Many STI prevention programs in schools and communities encourage people to make an explicit commitment to practice safe sex. This means encouraging people to identify their personal sexual goals and identify and implement the strategies that they will need to follow to achieve those goals in their own lives. The more public the commitment, and the more preparation that goes into the strategies, the more likely people are to follow through on their commitments.

6.3. Sexual Communication Skills

People, especially women, can reduce their vulnerability in sexual situations by communicating more effectively and assertively with their partners. First, this involves educating oneself about STIs, their risks, and what one needs to do to protect oneself. Next, this means initiating desired activities and refusing unwanted activities, and negotiation about condom use. Aside from condom use, a couple may negotiate such sexual matters as sexual exclusivity and getting HIV antibody testing together. Being direct and honest with potential partners about one's own health and risk status is key, as is getting relevant, honest information from one's potential partner about whether the partner has, for example, had a recent HIV test or has ever had unprotected anal sex, sex with commercial sex workers, or engaged in intravenous drug use. Unfortunately, even honest answers to direct questions may not be enough because many people are unaware of the risks associated with some behaviors and of their own disease status. Moreover, many surveys show that a significant minority of people—approximately 20–35%—lie about their sexual history or disease status to avoid rejection, further complicating negotiated safety.

The most important communication about sex takes place between people who are contemplating sexual relations, but open communication about sex should

occur in other forums and among other groups of people as well. Research suggests that parents can play a constructive role in affecting their children's sexual risk-taking. Mothers appear to be particularly critical in the development of responsible adolescent sexual behavior. Open communication in a nonjudgmental way between mothers and teens about sexual matters and parental monitoring of teen behavior predicts lower sexual risk-taking in adolescence.

6.4. Avoiding Known High-Risk Behaviors

Alcohol and other drugs do not cause STIs, but they are consistently associated with high-risk sexual behavior because they can be disinhibiting and interfere with a person's judgment and resolve. If one remains sober, one is more likely to use condoms and less likely to engage in high-risk sexual behaviors.

6.5. Getting Regular Medical Checkups

Early detection of STIs is critical to treating them effectively. Moreover, the presence of untreated STIs makes new infections, including HIV, more likely. Many STIs, such as syphilis and chlamydia, have few or no symptoms and, if untreated, can have serious consequences. Routine screenings—approximately twice a year—are highly recommended for sexually active people generally and for couples about to become sexually active. Many public health centers provide such services for free or at a low cost.

See Also the Following Articles

Homosexuality ■ Sexual Behavior ■ Sexual Behavior and Culture

Further Reading

Buss, D. M. (1989). Sex differences in human mate preferences. Evolutionary hypotheses tested in 37 cultures. *Behavioral and Brain Sciences*, *12*, 1–49.

Centers for Disease Control and Prevention. (2000, June 9). Youth risk behavior surveillance—United States, 1999. *Morbidity and Mortality Weekly Report*, *49*(SS05), 1–96.

Cooper, M. L., Shapiro, C. M., & Powers, A. M. (1998). Motivations for sex and risky sexual behavior among adolescents and young adults: A functional perspective. *Journal of Personality and Social Psychology*, *75*(6), 1528–1558.

Fine, M. (1988). Sexuality, schooling, and adolescent females: The missing discourse of desire. *Harvard Educational Review*, *58*, 29–53.

Fraser, A. M., Brockert, J. E., & Ward, R. H. (1995). Association of young maternal age with adverse reproductive outcomes. *New England Journal of Medicine*, *332*, 1113–1117.

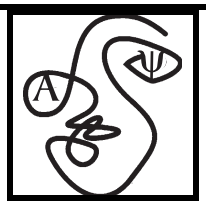
Kaiser Foundation. (2002). *HIV/AIDS policy fact sheet*. Washington, DC: Henry J. Kaiser Family Foundation.

Kelly, G. F. (2004). *Sexuality today: The human perspective* (7th ed.). Boston: McGraw-Hill.

McAnulty, R. D., & Burnette, M. M. (2004). *Exploring human sexuality: Making healthy decisions* (2nd ed.). Boston: Allyn & Bacon.

Rathus, S. A., Nevid, J. S., & Fichner Rathus, L. (2002). *Human sexuality in a world of diversity*. Boston: Allyn & Bacon.

Tolman, D. (2002). *Dilemmas of desire*. Cambridge, MA: Harvard University Press.



Robotherapy

Elena V. Libin
Institute of Robotic Psychology and Robotherapy, Cyber-Anthropology Research Incorporated, Chevy Chase, Maryland, USA

Alexander V. Libin
Georgetown University, Washington, DC, USA

1. Robotherapy: A New Era of Interactive, Nonpharmacological Interventions
2. A Concept of Artificial Partners
3. Robots as Therapeutic Tools: Early Experiments in the Field of Robotherapy
4. Person–Robot Communication as Complex Interactive System
Further Reading

GLOSSARY

artificial partner An interactive robotic or digital creature that provides positive sensorimotor, cognitive, or emotional stimulation.

nonpharmacological therapy The use of psychologically based or activity-based techniques for treating a person's problematic behaviors and negative emotional states.

robotherapy A framework of human–robotic creature interactions aimed at the reconstruction of a person's negative experiences through the development of coping strategies, mediated by technological tools, in order to provide a platform for building new positive life skills.

The use of artificial interactive creatures as therapeutic tools is a practical application that resulted from a newly developed theoretical and methodological field called robotic psychology and robotherapy. The innovative concept of robotherapy offers methodological and experimental justification for the use of nonpharmacological interventions based on stimulation, assistance, and rehabilitation techniques for people with cognitive impairments, physical disability, special needs, and/or psychological problems. Robotherapy has been defined as a framework of human–robot interactions aimed at the reconstruction of a person's negative experiences through the development of coping strategies mediated by technological tools in order to provide an emotional and cognitive platform for building new positive life skills. Field experiments have shown the effectiveness of robots as human companions and therapeutic agents.

1. ROBOTHERAPY: A NEW ERA OF INTERACTIVE, NONPHARMACOLOGICAL INTERVENTIONS

A great variety of therapeutic techniques and theories traditionally related to mental health can be reduced to a basic dichotomy, defined as pharmacological vs nonpharmacological approaches to treatment. The main effects of pharmacological interventions are based on drugs' impact on human physiology associated with biochemical processes. Nonpharmacological approaches for treating problems that a person experiences during the course of his or her life are aimed at psychological mechanisms that underlie behaviors that are disturbing to the person or others. Psychological effects produced

©2004 Elsevier Inc.
All rights reserved.

by nonpharmacological interventions influence biochemical and neurological patterns that form individual behaviors, states, and traits. Psychological therapy achieves the same effects as pharmacological methods, affecting the wide range of mechanisms that underlie psychophysiological manifestations of emotions and complex social problematic behaviors. Among psychologically oriented techniques, a new group of technology-mediated therapy is rapidly developing. The invention of the microchip brought new technological tools to the therapist: telemedicine, virtual reality, computer games, and robotic creatures. Four major approaches are distinguished in technology-mediated therapy:

1. Virtual reality techniques for treating anxiety, attention deficit disorder, fear of flying, and a variety of phobias based on the phenomenon of presence and immersion
2. Internet-based communication, including telehypnosis, distant psychotherapy, and therapy-oriented knowledge acquisition from electronic databases
3. Electronic games tailored for clinical application
4. Robotherapy based on a person's interactions with embodied agents, such as social and entertainment robots, with different levels of artificial intelligence and integrated sensory feedback

The use of artificial interactive creatures as therapeutic tools for medical and psychological practice is the subject of a new field of research called robotherapy. The first definition of robotherapy by Libin and Libin in 2002 describes it as a framework of human–robotic creature interactions aimed at the reconstruction of a person's negative experiences through the development of coping strategies that are mediated by technological tools in order to provide a platform for building new, positive life skills. Robotherapy deals with the analysis of person–robot communication, viewed as a complex interactive system, while emphasizing psychological evaluation, diagnosis, and prognosis of problem behaviors and principles of nonpharmacological treatment.

2. A CONCEPT OF ARTIFICIAL PARTNERS

A paradigmatic shift in human–computer interaction studies a range of areas, from human factors to personal robot assistance, and outlines a new class of robots with therapeutic potential. In 2001, Libin defined artificial partners as interactive robotic or digital creatures that provide positive sensorimotor, cognitive, or emotional stimulation for a human partner. The concept of artificial partner (ART) places the relationships between humans and robots into a psychological, rather than technological, context.

The class of interactive stimulation robots as human companions is distinguished by two characteristics that make it potentially valuable for sociopsychologically oriented research and practice. First, the design and behavioral configurations of interactive artificial creatures are tailored to the purpose of communicating with a human being on a "personal" level. The interactions between a person and a robot include tactile–kinesthetic, sensory, emotional, cognitive, and social–behavioral patterns. Second, ARTs imitate real-life (human- or animal-like) behavior and model motor, emotional, and cognitive behaviors normally experienced by living beings. These features form the artificial partner's functional status, which reflects a structure of a living world, making interactive agent an interesting and engaging communicating, gaming, or therapeutic partner for people of all ages, cultures, and life experiences. Numerous artificial creatures that are equipped with touch, audio, visual sensors, and different levels of robo-IQ exist on the market and in the laboratory. Anthropomorphic robot platforms Cog and Kismet, Doc Beardsley and Nursebot Pearl, humanoids AMI and ASIMO, emotional robotic cat NeCoRo, robotic dog AIBO and robot seal Paro, and automated dolls Amazing Amy and My Real Baby are the most vivid examples.

3. ROBOTS AS THERAPEUTIC TOOLS: EARLY EXPERIMENTS IN THE FIELD OF ROBOTHERAPY

The function of artificial partners is changing from a role as entertainment agents to include what were originally exclusively human occupations, such as education and therapy. Artificial creatures with high robo-IQs and integrated sensory, evaluating, and controlling systems perform certain functions in the therapy process, such as the following:

1. A diagnostic method for studying individual coping and defense machanisms
2. A tool for guided physical and mental stimulation
3. An interactive device for training and development of certain individual and group skills
4. A mediator in person-to-person communication

3.1. Coping with Negative Behavioral and Emotional States via Technology: Behavioral Therapy

Research with the lifelike robotic seal Paro, developed by Takanori Shibata in 1998, was aimed at studying the psychological, physiological, and social effects of robotic pet interactions with both children in the hospital setting and elderly persons in nursing homes. Paro is covered with synthetic fur, which makes interactions based on tactile–kinesthetic senses possible. Results showed that physical interactions mediated by the robot's tactile sensors positively influenced the subjective evaluation of the artificial creature's appearance. Analysis of attitudes toward Paro revealed two general factors that characterized the individual experience of communicating with the robot companion: The first factor was related to the unique tactile properties of the artificial creature, and the second factor was responsible for its communicative value. Analysis of psychophysiological and psychological indicators in the context of therapeutic effects produced by the human–robot communication showed that the texture and softness of Paro's fur, in combination with its empathic appeal, had a positive therapeutic effect on the mood and associated state of comfort in interacting with persons of different ages, genders, and medical and cognitive status. It is also promising for future research that the 2003 *Guinness Book of World Records* considered the Paro to be "the most therapeutic robot" ever designed.

3.2. Robots for Personal Assistance: Psychological and Physical Rehabilitation

Robotherapy studies have been conducted on diverse clinical populations, such as children and elderly with special needs, physically and cognitively impaired persons, and people with a wide range of psychological problems. In 1967, the founder of cybernetics, Norbert Wiener, noted that a major function of sophisticated communicative machines (as well as higher living organisms) is the ability to resist entropy by modifying their behavioral configurations on the basis of past experiences and by using sensitive feedback. Robotic creatures as therapeutic agents adjust to individual preferences and needs by collecting information associated with various modalities of the person involved with the treatment.

Allison Druin's team at the University of Maryland's Human Computer Interaction Laboratory developed robotic PETS or Personal Electronic Teller of Stories. Children customize the robot's appearance using different animal parts (dog paws, wings, horns, etc.) so that the robot has a personal representation. Children then create stories to express how they feel and program the PETS robot to express these emotions through movement and sound. This method of communication gives children with special challenges a sense of control and fulfillment. Through interactions with the robot, children overcome their own limitations of emotional and verbal expression. A storytelling robot has the potential to provide effective therapy for emotionally challenged children (e.g., those with autism).

Gestural interface technology is based on recognition of human nonverbal behavior and associated emotions, and it is the core of the tactical mobile robot called JesterBot. When attached to a person in a wheelchair, JesterBot collects signals from tiny muscle movements and transforms them into meaningful verbal and physical interactive patterns. The robot modifies its own behavior by monitoring the vital signals and retrieving records about the person's sensorimotor profile.

CosmoBot, invented by Corinna Lathan's team at AnthroTronix, Incorporated, is designed to be a companion for children working on developmental goals associated with physical, occupational, or speech/language therapy or on special education goals. An interactive rehabilitation robot detects the child's movements and/or voice commands, transforming them into command signals to make the robot move and/or talk. This rehabilitation aid is programmable, allowing the child to interact with CosmoBot in the context of new games and stories. Furthermore, the robot's data-collection capabilities allow therapists and educators to objectively track the child's progress with regard to developmental goals.

Robots assist mentally and physically challenged persons by facilitating support for their daily living activities through the enhancement of their sense of control. In 1993, Lancioni *et al.* developed a robot-assisted occupational program that provides choice opportunities for people with severe retardation and multiple handicaps. Another approach employed a group of robots with different design and functionality for a group of neuropsychologically impaired persons in order to compensate for their limited cognitive and physical resources. Individuals' comprehension abilities, physical needs, and communicational styles were matched to the robots' capacity to provide visual, auditory, and motor support and communicative stimulation. In 2002–2004, Libin

and Libin studied the impact of robotic-based entertainment systems on positive, emotional experiences and disruptive behaviors in persons with limited physical, sensorimotor, and cognitive resources. Libin and Cohen-Mansfield showed that for elderly people with dementia, interactions with the robotic pets resulted in decreased agitation and increases in positive emotions.

Interactive robots are capable of promoting sensorimotor activities, triggering cognitive and emotional stimulation, and providing an individual with psychological benefits (i.e., a sense of control, independence, and self-efficacy). The effectiveness of the robotic devices implemented during the course of therapy depends on a user-friendly design, a software program that provides choice-making behavior, and the feedback system with verbal and nonverbal reinforcement.

4. PERSON–ROBOT COMMUNICATION AS COMPLEX INTERACTIVE SYSTEM

The theoretical framework of robotherapy, in which an artificial creature communicates with a person for assistance or therapeutic purposes, stems from the concept of complex interactive systems, which is based on three principles. First, the structure of person–robot communication can be viewed as a constellation of associated exchanges between patterns related to different modalities. The robot as an embodied therapeutic function coexists with an interacting person in the same given space–time continuum, becoming an interactive agent. This quality of an artificial creature is vitally important for people with limited resources (i.e., physical, cognitive, or emotional impairment) or persons who strive toward further development through mastery of certain skills. Second, because an interactive system behaves equifinally ("the same goal can be reached from different initial conditions or in different ways," as von Bertalanffy noted in 1966), robots with various levels of artificial intelligence and types of sensory feedback can benefit people with different combinations of skills and abilities as an effective tool in a very broad context of individual situations. Lastly, a systematic approach concentrates on the issue of matching a robot's behavioral configuration with individual profiles. The components of this interactive system are summarized in Fig. 1. The model was developed using available data derived from studies on social robotics, computational

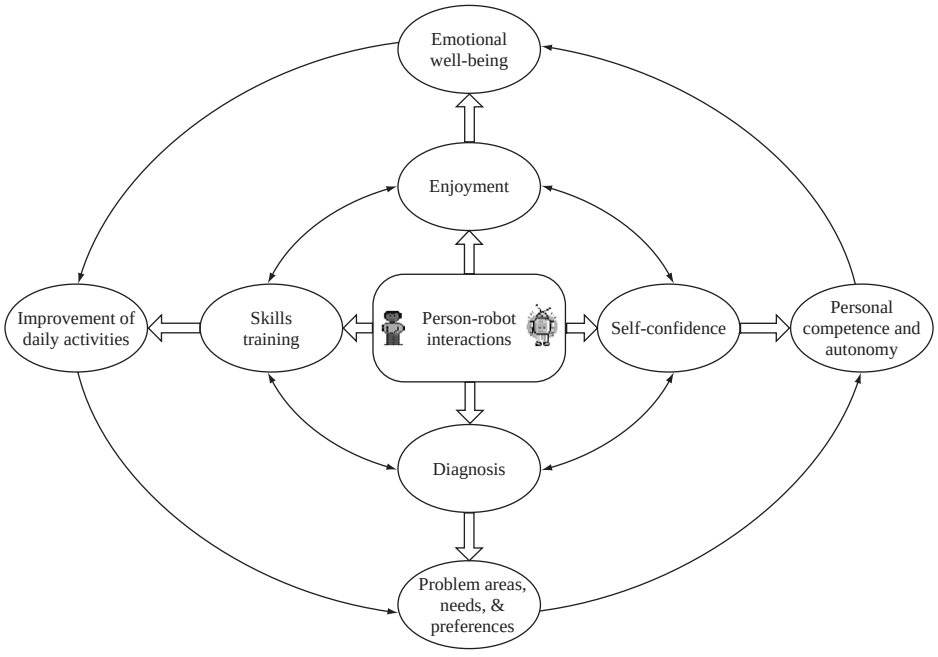


FIGURE 1 Multidimensional model of a person–robot communication as a complex interactive system. The following are the correlations between system parameters and types of interactions: skills training—sensorimotor stimulation, attention, and memory; diagnosis—identification of problem areas, individual needs, and preferences; self-confidence—increasing personal competence and autonomy; and enjoyment—enhancing emotional well-being.

entertainment, modern information technology, and cybertherapy.

Two loops underlying the person–robot communication describe elements of a complex interactive system. The internal loop is formed by the patterns that underlie skills training, problem diagnosis based on monitoring vital signals, enjoyment resulting from emotional stimulation, and self-confidence developed through the sense of control. The external loop embraces the results of the internal loop functioning, such as improvement of professional or daily activities due to sensorimotor and cognitive training; positive well-being; personal competence and autonomy; and mental and physical profile highlights tailored to one's problem areas, needs, and preferences. The internal loop produces such integrative system effects as compensation and optimization in a person's functioning, whereas the external loop yields effectiveness and adequacy of interactions.

The effectiveness of robotherapy is influenced by (i) one's past experiences, current needs, and individual preferences and (ii) an artificial partner's nontransitive physical features and behavioral configurations defined through the intensity of stimulations and responses. At the same time, the effectiveness of robotherapy is associated with the productive compatibility between human beings and their artificial partners, which is achieved through systematic study of individual needs and preferences, user-friendly design and likeable appearance of the robotic creatures, adjustable behavioral configuration of a robot, and multilevel (tactile, sensorimotor, cognitive, emotional, and social) person–robot interactions based on the open-loop principle.

As a new way of helping people, robotherapy is aimed at the improvement of personal well-being through the development of various coping skills mediated by technological tools. The use of robots in therapy poses a basic moral dilemma regarding whether robotic assisted care should be a substitute for human care. The implementation of smart interactive technology for the treatment of people raises questions about the responsibility of the therapist for the quality of care. Robotherapy employs a humanistic approach in psychological consulting and therapy, according to which a psychological culture of the therapist defines the therapeutic value and the proper use of technological innovations.

See Also the Following Articles

Coping ■ Robotic Psychology

Further Reading

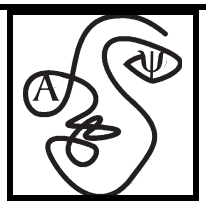
Fukuyama, F. (2003). *Our posthuman future*. New York: Picador.

Libin, A., & Libin, E. (2004). Person–robot interactions from the robopsychologist's point of view: The robotic psychology and robotherapy approach. In *Person–Robot Interactions for Psychological Enrichment*, Special Issue. Los Alamitos: IEEE.

Ludwig, L. (2001). *Robo sapiens*. Cambridge, MA: MIT Press.

Mahony, M. (2000). *Cyborg: The man machine*. Bradford: Bradford Press.

von Bertalanffy, L. (1967). *Robots, men, and minds: Psychology in the modern world*. New York: George Braziller.



Robotic Psychology

Alexander V. Libin
Georgetown University, Washington, DC, USA

Elena V. Libin
Institute of Robotic Psychology and Robotherapy, Cyber-Anthropology Research Incorporated, Chevy Chase, Maryland, USA

1. Robotic Psychology: On the Coexistence of Humans and Robots
2. Psychology of Person–Robot Interactions
3. Companionship between People and Robots
Further Reading

GLOSSARY

robo-IQ The complexity of the software-based artificial intelligence related to an adequate performance of an uncertain or multiple-choice task.

robot An automat with programmed behavior designed for the purposes of substituting for certain human activity or to satisfy a person's physical, psychological, or social needs.

robotic psychology A study of compatibility between people and artificial creatures on different levels: sensorimotor, emotional, cognitive, and social.

Robotic psychology is defined as an interdisciplinary field of research and practice that focuses on the compatibility between people and artificial creatures on different levels: sensorimotor, emotional, cognitive, and social. Robotic psychology studies the psychological significance of robots' behavior and its intertwining with elements of physical and social environments. Different classes of robots are aimed at fulfilling different human needs. Artificial creatures have their own distinct individuality, which manifests itself in the robot's design and behavioral configuration, in the same way that people and other living beings differ from each other in various parameters, such as weight and height, behavioral reactions and character, emotions and cognition, abilities, and coping strategies. Analyses of person–robot interactions account for the essentials of both humans' behavior and that of their artificial partners.

1. ROBOTIC PSYCHOLOGY: ON THE COEXISTENCE OF HUMANS AND ROBOTS

Interactions with robotic creatures of different kinds are rapidly becoming part of people's daily lives. Satisfying various human needs, robots with artificial intelligence and sensitive feedback perform industrial and military functions, research and medical procedures, and assist in educational and therapeutic tasks. A study on human–robot coexistence based on innovative robotic psychology concept now encompasses cybernetics and humanities, engineering and biology, and information sciences and aesthetics. An analysis of robotic beings and their behavior from the psychological standpoint opens new perspectives for theoretical and practical applications in both technological and social sciences. Robotic psychology as a new

©2004 Elsevier Inc.
All rights reserved.

research area employs a systematic approach in studying psychophysiological, psychological, and social aspects of person–robot communication.

The search for an artificial being as a human companion dates back to ancient mythology and creatures such as Golem in medieval literature. However, the contemporary understanding of an automatic creature began when Czech writer Karel Capek coined the term "robot" (meaning a worker or a servant) in 1921 and Russian-born U.S. author Isaac Asimov developed the concept of robotics in the 1940s to describe the field of designing and using automats for human purposes. The term robot can be defined nowadays as an automat with programmed behavior aimed at substituting for certain human activities or to satisfy a person's physical, psychological, or social needs.

1.1. Classification of Artificial Creatures in Robotic Psychology

Due to the diversity of human activities, a variety of robotic creatures are required to help in the performance of these activities. From the perspective of human needs, all kinds of interactions between people and robots deserve psychological attention. Contemporary primary classification of artificial creatures is based on a robot's relation to human activity, a robot's behavioral configuration defined by degrees of freedom, robo-IQ (complexity of software-based artificial intelligence related to an adequate performance of an uncertain or multiple-choice task), and the robot's physical properties manifested in the hardware architecture and appearance (Table I).

Robotic psychology emphasizes three major modes of a robot's relation to human needs: substitute for human activity (e.g., rescue robots), extension of human activity (e.g., medical robots), and stimulation of human activity (e.g., social robots and therapeutic robots).

2. PSYCHOLOGY OF PERSON–ROBOT INTERACTIONS

Rodney Brooks, a scientist from the Massachusetts Institute of Technology (MIT), developed a concept of robotic beings. He argued that robots' behavior must be based on the principle of dynamic interactions with changing physical and social environments. A robotic being is perceived as an embedded part of human relations with the world, as one of the components of the interactive process but not the separate system. To be relevant to the unpredictable psychological reality of its human partner, the organization of a robotic creature's behavior needs to be built with the motor and sensory loops running in parallel rather than employ hierarchically structured functioning.

2.1. Modeling the Human Mind

Robots' capacity to adjust to the requirements of the human world stems from the concept of artificial intelligence, or a set of programs that control the automatic actions of machines. John McCarthy, who coined the term artificial intelligence (AI) in 1956, was an enthusiast of the psychologization of computers. This position reflects an idea that there is no difference between human reasoning and computer programming. A different approach expressed by philosopher John Searl states that although the human mind has the ability to compute, it is completely different in its nature from AI. Whereas robotic actions result from embedded programs, human behavior is determined by subjective needs and preferences, likes and dislikes, and intentionality. That is why it is such a challenge to build an artificial creature capable of interactions relevant to the complexity of human relationships. It is of no coincidence that the competence of a robot's AI is tested by the human expert implementing the Turing test, based on the robot's ability to answer random questions in the same manner as people do.

On the other hand, it is considered an honor to challenge an artificial partner in fair competition, such as in a game of chess. Grossmeister Gary Kasparov, along with the rest of the world, regarded IBM's Big Blue robot as a serious gaming partner in a person–robot chess competition. Sophisticated artificial intellect requires an adequate vehicle. In 2003, Hugo de Garis developed an artificial brain with 37.7 million artificial neurons that will allow a robo-kitten designed by Michael Korkin to perform an animal-like intelligent behavior based on retrieving sensory experience from the different parts of the environment.

2.2. Modeling Emotions

The idea of modeling intellectual functions of the human mind has existed throughout the history of robots. The reproduction of human emotions did not attract investors' attention as much as the built-in cognitive component of artificial creatures. Nevertheless, the new millennium was marked by the paradigmatic shift in robot design from mechanocentric to human-oriented

TABLE I
Classification of Artificial Creatures in Robotic Psychology

	Human need	*Behavioral configuration*	*Physical appearance*
Industrial robots	Perform hard labor and hazardous work		
Research robots	Provide extension of human sensorimotor capacities	Defined by a situated activity that substitutes for human behavior	Machine-like appearance with emphasis on peripheral devices
Military and rescue robots	Activity in life-threatening situations		
Medical robots	Performance of fine motor operations on the human body		
Recreational robots	Trigger positive behaviors through entertainment	Aimed at human feelings and emotions	Animated form of existing and non existing beings or objects
Interactive stimulation robots			
Social robots	Communication	Based on human facial expressions and complex gestures with social meaning	Anthropomorphized (humanoid) appearance
Robots with therapeutic potential	Therapy and treatment of negative emotional states and behaviors	Models basic emotional states and social behaviors embodied in lifelike form	Animal-like or human-like appearance

principles, which transform the very purpose of artificial creatures' existence, their architecture, and their physical appearance.

A new concept of an artificial emotional creature proposed an original lifelike design and behavioral configuration relevant for natural human communication. The robotic cats Tama, designed in 1998 by Toshihiro Tashima, and NeCoRo, created by Takanori Shibata and colleagues, employ fuzzy logic in the organization of AI, which makes the robotic creatures' behavior "unpredictably" engaging and emotionally attractive for humans. A distinct feature of emotional robot pets is the use of tactile sensors covered with "skin," a synthetic fur. The combination of lifelike behavior and tactile stimulation creates an appeal that is missing in many other robotic models. In 2002, Libin and colleagues studied the influence of past experiences with real animals and technology on a person's emotional communication with a robotic creature, the Omron Corporation's robotic cat NeCoRo. The intensity and direction of emotional responses triggered in a human partner by the robot were associated with the human's relevant past experiences. People perceived the robots as part of their living world.

2.3. Modeling Social Behavior

Regardless of shape (human or animal), the robot's most difficult task is to model people's social behavior with such cues as eye contact, gestures, and emotional answers recognizable in the social context. Social robots model the natural process of communication and are able to learn through social interactions. Braezeal and colleagues applied a developmental psychology approach to the design of the autonomous robot Kismet, build at the MIT media lab in 1998. The learning processes of robots are similar to those of children and are described by Lev Vygotsky's concept of the proximity zone; that is, the complexity of the stimuli increases when a learning agent reaches a certain level of competency in performing the task. In the robot's process of social learning, this condition is achieved by maintaining a level of interactivity that is neither overwhelming nor understimulating.

2.4. Modeling Teaching Behavior

During the education and learning process, both the levels of robo-IQ or communicational ability and the robot's anthropomorphic characteristics are important. Children ages 3–5 years are more motivated to learn from a personified mobile robot with a smiling face and verbal positive reinforcement than from an immobile and nonverbal machine-like device.

Analysis of children's needs and preferences resulted in the creation of a realistic robotic doll called My Real Baby, which enhances a child's communication experience via 15 human-like emotions of different intensity: facial expressiveness, soft and flexible skin, and a rich vocabulary of sounds and words. This doll, which is in the class of interactive stimulation robots, allows a child to develop educational interactive play experiences triggered by his or her imagination and creativity through natural responses expressed in emotionally appropriate ways.

A variety of humanoid robots, such as Kismet and Cog, Doc Beardsley and Nursebot, and AMI and PINO, are able to provide warmth and friendliness for their human partners. Another option considers animals, especially pets, as prototypes for the robotic creature design. Among the most interesting examples are the robot seal Paro and horse Karakuri, dog AIBO and cat NeCoRo, and robo-puppy Techno and robo-kitten Robokonenko. Robotic pets' resemblance to real animals and associated human experience allowed Peter Kahn and colleagues in 2002 to study the influence of communication with robotic dog AIBO on children's cognitive, social, and moral developmental processes. Children's conceptualization of the robotic creatures' behavior was similar to their perception and reasoning based on experience with real pets.

3. COMPANIONSHIP BETWEEN PEOPLE AND ROBOTS

A core aspect of the concept of companionship between people and robots is their compatibility. Compatibility in the robotic psychology context depends on how humans perceive their robotic companions and how they feel about them emotionally. In a study conducted by the Sakamoto Laboratory of Ochanomizu University in collaboration with the NEC company, three robotic creatures—dog AIBO and humanoids ASIMO and PaPeRo—were compared with each other as well as with animals, humans, and inanimate objects. People's perceptions of the robots differed from their perceptions of animals, other humans, and inanimate objects. The appearance and shape of the robots influenced the level of comfort people experience around technological devices, especially robotic creatures.

In 2002, Libin, Libin, and colleagues performed a cross-cultural study of the United States and Japan that investigated biological, psychological, and cultural determinants of people's communication with the robotic cat NeCoRo. A culture was defined as a system of created social meanings and artifacts that have symbolic value and reflect the logic of human mentality derived from norms shared by a given community. A specially designed scale evaluated verbal and nonverbal, emotional, and tactile–kinesthetic patterns of communication. For age and gender, preferences regarding the robot's features and interactive patterns were similar in both cultures. Older people liked the robotic pet more, whereas men enjoyed the cat's triggering behavior more than women. The influence of a person's past experiences with live pets and modern technology on the type of person–robot communication was universal. Differences between cultures were found for the participant's evaluation of the robotic creature's appearance and character and the intensity of tactile interactions.

Robotic psychology research incorporates the elements of training, education, entertainment, and therapy into the person–robot communicative process, which leads to the development of an enhanced constructive model of interactive technology.

See Also the Following Articles

Emotion ■ Learning ■ Robotherapy ■ Social Skills Training ■ Teaching Effectiveness

Further Reading

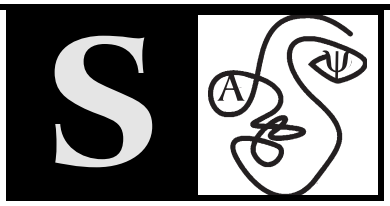
Arkin, R. (1997). *Behavior-based robotics*. Cambridge, MA: MIT Press.

Braezeal, C. (2003). *Designing sociable robots*. Cambridge, MA: MIT Press.

Brooks, R. (2002). *Flesh and machines: How robots will change us*. Cambridge, MA: MIT Press.

Libin, A., & Libin, E. (2003). Robotherapy and robotic psychology: New diagnostic tool for measuring person–robot communication. *Cyberpsychology and Behavior*, 6(4), 369–374.

Moravec, H. (2001). *Robot: Mere machine to transcendent mind*. Cambridge, MA: MIT Press.



Safety and Risks: Errors and Accidents in Different Occupations

Doris Fay and Patrick Tissington

Aston University, Birmingham, United Kingdom

1. Introduction
2. Errors: From the "Freudian Slip" to Taxonomies of Errors
3. Strategies of Error Prevention
4. Antecedents of Accidents: The Search for Individual Difference Variables
5. From Engineering Control Solutions to Safety Behavior
6. Recent Approaches to Safety: The Role of Management and Organizational Climate
7. Safety-Enhancing Interventions
8. Summary

Further Reading

GLOSSARY

accident An unexpected event, damaging person(s) and/or a system, involving the potential to impair the goal accomplishment.

error Intentional action that does not result in the attainment of the goal; prevention of the error must have been under the control of the person(s), thereby excluding an "act of God" as the reason for the nonattainment.

human factors The study of humans and their interaction with machines and the environment.

safety climate Employees' perception of the organization's stance on compliance with safety-based rules and procedures and their perception of the priority that management accords to occupational safety.

system design The deliberate application of science to the planning and specification of the operation of machines by humans.

unsafe behaviors Behaviors that are not in accordance with safety rules and procedures; they enhance the likelihood of accidents.

Errors and accidents are a severe threat to the well-being and performance of the individual worker, of the social and technical work system, and of the wider public and environment. Therefore, understanding factors and interventions that reduce errors and enhance safety is of interest for multiple stakeholders. This article reviews the diverse theoretical and practical approaches to errors and accidents. It introduces prominent error taxonomies and strategies of error reduction such as appropriate system design. Contemporary models of accidents acknowledge that they are the result of a multifactorial process in which errors can play one part. The article highlights factors such as quality of system design, individuals' safety behaviors, behaviors of management, and the organization's safety climate. It discusses interventions that relate to the various factors involved.

©2004 Elsevier Inc.
All rights reserved.

1. INTRODUCTION

"A man who makes no mistakes does not usually make anything." This quotation, from Bishop William Conor Magee, makes the case that errors are a natural and unavoidable component of human activity, and some theories of individual and organizational development take a positive stance to errors. Such theories take the view that errors contribute to the development of human knowledge and skills in that they represent opportunities for organizational learning. However, errors can be seen as a double-edged sword. In occupational settings, errors predominantly have negative connotations because they are associated with reduced performance and quality. For example, in manufacturing, errors lead to higher costs through time spent on rectification or product recalls. In health care, errors can involve severe threats to the safety of patients and care providers. Furthermore, errors are frequently causal factors in accidents. For example, errors in power plants or chemical production can cause accidents that have significant consequences to the health and well-being of the people immediately involved, to the wider public, and to the environment.

On March 6, 1987, the ferry *Herald of Free Enterprise* capsized and sank outside the Belgian port of Zeebrugge, resulting in the loss of 188 lives. The cause of the accident became clear almost immediately. The ship had set sail with its large bow car-loading doors open, allowing water to cascade into the car deck as the ferry increased speed. The subsequent inquiry found that the root cause of the accident was a series of errors by individual operators, designers, and the organization's senior management. Essentially, the accident was caused by a junior member of the crew failing to operate a series of levers that would have closed the watertight doors. However, there was some doubt as to whose responsibility this was, given that the captain did not know that the doors were open. Many requests had been made by officers to add a warning light on the bridge to indicate when the doors were open, but these requests had gone unheeded by the company's managers. Policies on status reporting were confusing but were founded on a principle of "negative reporting only," so there was no emphasis on checking. Clearly, this was a tragedy on a large scale and was of great interest to researchers in accidents and errors due to the multiply stranded antecedents of the accident. This one incident graphically demonstrates the potentially disastrous end result of errors and shows that the study of errors in organizational contexts is a vital area of study and must include both rigor and practical application. The scientific approach to accidents and their causes is to develop models of accident causation.

Contemporary models of accidents acknowledge that accidents are the result of a multifactorial process in which errors play a part. Furthermore, the display of unsafe behaviors that, in turn, influence the likelihood of an accident is given a pivotal role. Models suggest that diverse factors such as managerial strategies, design of the system, equipment, design of the entire job, characteristics of the worker(s), and social environment influence unsafe behaviors. Figure 1 illustrates a number of factors that are relevant to accidents. The costs involved in errors and accidents are social, economic, and psychological ones. They include individuals' pain and suffering, insurance settlements, indirect costs such as investigation time and downtime, and replacement of worker(s) and/or equipment. Therefore, understanding factors that assist in the prevention of errors and accidents is critical. In what follows, the article describes current approaches to errors, discusses factors contributing to safety, and then shows how they are used for safety interventions.

2. ERRORS: FROM THE "FREUDIAN SLIP" TO TAXONOMIES OF ERRORS

The study of error, in terms of its prevention, has been the subject of research by psychologists for more than 100 years. An early attempt to explain error in psychological terms was made by Freud, who viewed error as being indicative of an underlying need or some other purpose—the so called "Freudian slip." Consider the case of a train driver passing a signal that was set at "danger"—an error worthy of study for obvious reasons. Tracing this error to some kind of psychosexual origin in the driver has little currency today. Many other approaches to error have focused on the person making the error and attributed error to carelessness or inattention of the individual concerned—in other words, attributing blame to the driver of the train in this example, or to the operator of the watertight bow door in the example of the Zeebrugge ferry, and possibly reprimanding him or her. Modern researchers question the value of this approach because it tends to correct a very specific error made by an individual on a specific occasion and, therefore, is unlikely to aid in the prevention of future errors. Current thinking favors the development of

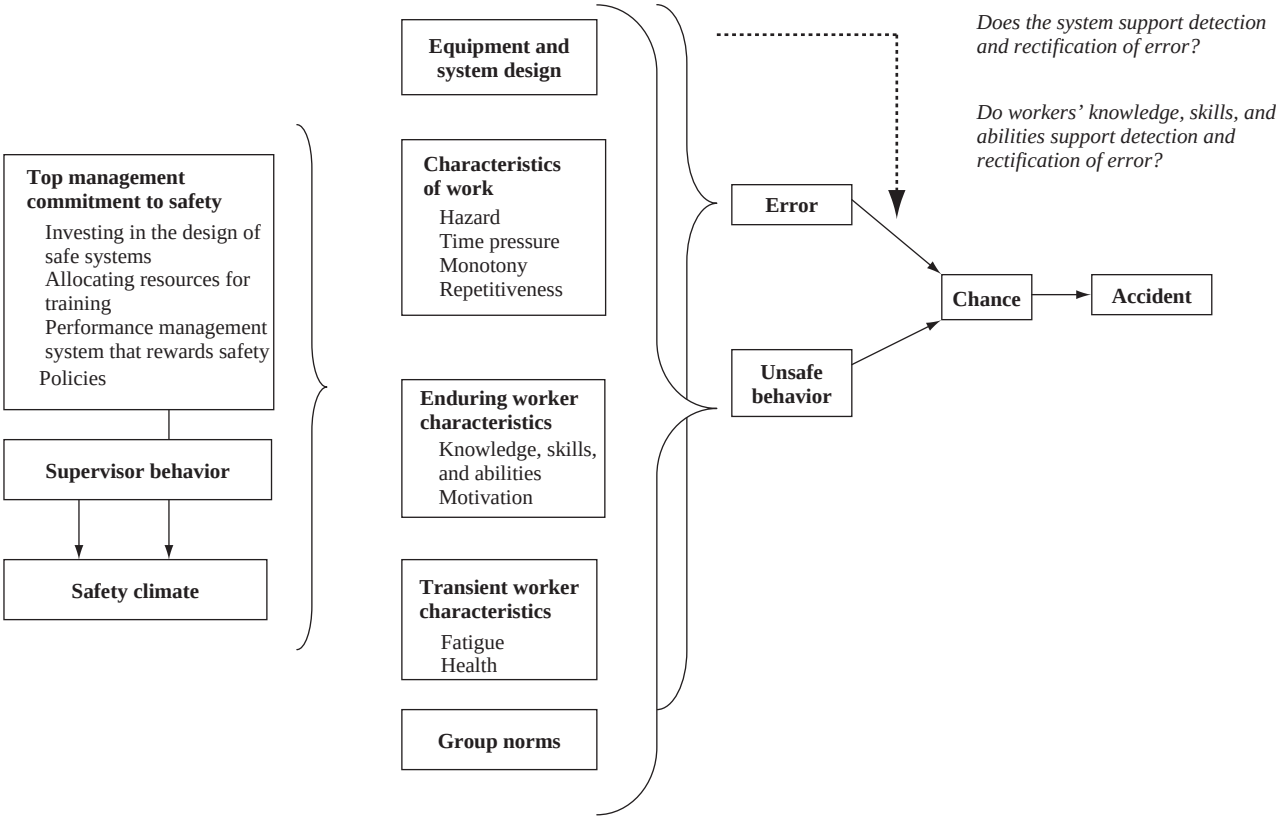


FIGURE 1 Factors contributing to errors and accidents. Adapted from Sanders and McCormick (1993).

domain-specific taxonomies that may be used as frameworks to investigate the errors preceding accidents.

In 1990, Reason presented a widely used definition of error that conceptualizes error as "a generic term to encompass all those occasions in which a planned sequence of mental or physical activities fails to achieve its intended outcome and when these failures cannot be attributed to the intervention of some chance agency." The study of error has evolved to consist largely of highly domain-specific research that leads to highly domain-specific taxonomies. A large number of taxonomies with both overlapping and domain-specific features are available. This section reviews the taxonomies and approaches of Norman, Rasmussen, and Reason as well as the work of Zapf and colleagues. These models have been selected not only on the grounds of their prominence in the literature and their perceived application in a number of workplace situations but also because they are relatively distinct.

In 1981, Norman generated a three-category taxonomy of errors: slips, mistakes, and violations. A slip is a good plan that is poorly implemented, a mistake is an incorrect intention, and a violation is deliberate (for whatever reason). In the case of the train driver, a slip would be the driver attempting to stop the train but pressing the wrong lever, a mistake would be the driver not recognizing the signal as being set at danger, and a violation would be the driver deciding to ignore the signal (perhaps thinking the signal was faulty). It can be seen immediately that the framework assists in tracing the source of the error and helps the organization to construct a remedy.

In 1983, Rasmussen developed a framework that approaches errors by looking at the cognitive processes behind task performance and, therefore, the source of error. His cognitive control taxonomy describes three levels of cognition involved in task performance: skill, rule, and knowledge. Task performance ranges from sensorimotor, automatically performed actions to actions that are conscious and based on directed deliberation. Behavior is skill based when stored patterns of information are automatically recalled from past experience and enacted. An experienced typist moving from a German-language keyboard to an English-language

keyboard—where the "y" and "z" keys are interchanged—is likely to make skill-based typos. Rule-based actions are directed by rules copied from procedures that are stored in memory in the "if (condition) then (action)" form. A typical error that would be considered rule based is when an individual misinterprets the situation and chooses the wrong rule. The key difference between skill-based and rule-based behaviors is the extent to which conscious attention is directed at the task; skills are automatic, whereas rules are applied mindfully. In entirely unfamiliar situations where skills and rules do not apply, cognition moves to the knowledge level at which operation is based on one's study of the situation and the goal of the individual.

The cognitive control model has its origins in field research carried out in many domains, including the control rooms of nuclear power plants, and is a general model of cognition applicable in these domains. Therefore, the train driver not pushing the brakes hard enough would be a skill error, whereas the train driver not realizing that the signal (e.g., a red light) indicated danger would be a rule error. In the case where the driver had insufficient or inappropriate training, and so was uncertain whether or not to stop, a knowledge-based error would have occurred. In the example of the Zeebrugge ferry disaster, it would appear that errors occurred at the rule level (i.e., the rules were ambiguous). But when it was noticed that the bow doors were open, the personnel concerned also made an error at the knowledge level (i.e., they did not understand the gravity of the situation and take action to close the doors or alert the bridge).

In a further development, Reason combined his model with that of Rasmussen, proposing that an error classified as a mistake could be either rule based (i.e., using the wrong rule) or knowledge based (i.e., not knowing what the rule was) and providing a finer definition of the source of error.

In 1992, Zapf and colleagues developed an error taxonomy that also considers the different steps of the action process. An action encompasses the steps of developing a goal, collecting information on the environment (e.g., via attending signals), developing a plan and executing it, monitoring its execution, and processing feedback on the progress toward or achievement of the goal. Different levels of regulations control the action process. Similar to Rasmussen's model, levels of regulation range from skill-based regulation producing automatized behaviors to intellectual, knowledge-based actions. A fourth level of regulation represents the use of heuristics. This taxonomy crosses the steps of the action sequence with the four levels of regulation. This results in error categories that help to track the source of an error in terms of failure of memory, wrong movements, poor judgments, or other factors.

3. STRATEGIES OF ERROR PREVENTION

Selection, training, and design are the generic strategies used to prevent errors. Although selecting individuals with the relevant skills and abilities is certainly important, it is limited where jobs change and, hence, workers are subjected to new demands on their skills. Therefore, training is used to complement the former strategy. Finally, system design strategies that recognize human factors (e.g., the limits of memory capacity and of perceptual speed) can reduce the likelihood of errors. System design approaches range from attempting to make it impossible to commit an error (e.g., the train driver could not override a red signal because electricity for the train would be automatically cut when a signal was set to red) to making the system fail-safe. The latter relates to reducing the consequences of an error. In the case of the train driver, the error consequence could be a crash with another train. This consequence could be reduced by brakes being automatically activated when another object is detected on the track as well as by installing huge bumpers to buffer the hit.

The idea of fail-safe systems was taken further by Reason, who took the pragmatic view that in practice it is impossible to eliminate error completely. Therefore, the role of the organization is to protect against the consequences of error. In 1990, Reason conceptualized protection against error as a series of defenses. When all defenses fail in a particular way, an accident occurs. This is termed the "Swiss cheese" model because it conceptualizes errors as being like holes in a piece of Swiss cheese, whereby an accident occurs when all of the holes line up. This model is not of particular value to the analysis of incidents but is certainly the way in which many practitioners would consider the analysis of systems protecting against error and checking whether the "holes" in the "cheese" are not generally aligned. In the case of the train driver, one might consider whether the introduction of a system in which the driver could not override the signal might provide additional protection. One also might consider whether additional protection could be provided by adding further signals in particular stretches of track so that if one signal was passed, other signals might be used as a further indication to the driver.

Because it is impossible to design systems in which errors are entirely impossible, it is important that systems be transparent and allow the operator or user to intervene. System transparency helps to quickly detect an error so that it can be corrected before the negative error consequences unfold fully. Furthermore, the system needs to allow the operator to intervene so as to make rectification possible.

The development of computer programs for office work during the past two decades provides good examples of how the strategies have been implemented. Many programs that are typically used for office work follow the "what you see is what you get" strategy, meaning that whatever the user does is reflected in changes on the monitor (cf. Fig. 1). The systems have become more fail-safe. When the user does a potentially critical step (e.g., deleting a file), the program will ask for verification of the step. The program will also allow errors to be rectified given that "undo" options are designed into the systems.

After an accident, the organization frequently seeks to trace an error to a human source and then punish the individual. The use of the models described in this section will assist in addressing the underlying organizational causes of accidents by informing the choice of prevention strategy. This potentially leads to more lasting solutions.

4. ANTECEDENTS OF ACCIDENTS: THE SEARCH FOR INDIVIDUAL DIFFERENCE VARIABLES

An accident is defined as an unexpected event that results in injury to one or several persons and/or damage to a system and that potentially impairs the accomplishment of the system's goals. One influential early approach to occupational safety goes back to the presumed discovery of the accident-prone person. The observation that a small minority of workers seemed to suffer the majority of accidents led to the conclusion that some individuals are more likely to have accident than are others. This thinking, which located the cause of an accident within the individual, was very popular for some time. Apart from its intuitive appeal, it was attractive to organizations because it justified a shift away from designing safety into the work environment to selecting the right people for the job—a less costly accident prevention strategy. Currently, however, most researchers agree that there is no convincing evidence of accident proneness. Data showing that some individuals have more accidents than do others are explained in terms of differences in the degree to which individuals are exposed to hazards in the first place. Furthermore, wherever the law of probability operates, some individuals will have more than one accident by chance.

Related research has tried to identify personality variables that would predict the occurrence of accidents. Two recent pieces of work, a 1998 review by Lawton and Parker and a 2002 meta-analysis by Salgado, integrated the numerous individual pieces of research but could not produce a conclusive picture. Furthermore, this field is dominated by cross-sectional approaches in which many studies use companies' accident records and relate these to current expression of individual difference variables. Causal conclusions are likely to be flawed because the magnitude of some individual difference variables will reflect having had accidents rather than predicting accidents.

The only result on person variables that emerges consistently across different industries is the association between accidents and age. Younger people appear to have more accidents (an effect well known from vehicle driver accidents), and this proneness decreases with age. However, some findings indicate another increase in accidents after 50 years of age. In addition, accident severity, as reflected in degree of permanent disability or fatal accidents, seems to increase with age.

Despite the inconclusive result regarding the role of personality in accidents, this approach should not be dismissed entirely. First, it is possible that individual difference variables that contribute to successful coping with a workplace hazard differ from job to job. Whereas a high resistance to monotony might be important for some occupations, a high reaction time might be critical for others. Second, focusing more on personality variables that are related to the behaviors relevant for safety might be fruitful. For example, a study in the chemical industry found a relationship between psychological maladjustment and accidents. The extent to which individuals adhere to safety rules might be the link between the personality variable (maladjustment) and accidents. Finally, patterns or combinations of specific individual difference variables could deliver new insights.

5. FROM ENGINEERING CONTROL SOLUTIONS TO SAFETY BEHAVIOR

The engineering approach to safety represents an alternative to the previously described person-centered

perspective. The engineering approach seeks to design safety into the work system by acknowledging human factor variables. Numerous technical improvements have increased the safety of the physical working environment, tools, equipment, and machines. The effectiveness of the engineering approach is limited where it depends on whether the individual decides to make use of it. This is related to the three levels of prevention employed in this approach. The first level of prevention is to design a system such that the individual cannot engage in an action that could lead to an accident. For example, for some operations that involve the use of hazardous chemicals in the chemical industry, it is technically possible to design the system such that the worker cannot ever come into direct contact with the chemicals. In other words, the hazard has been designed out of the system. If this is not possible, engineering safeguards target the second level of prevention. This is done, for example, by installing a safety screen in front of a galvanic bath.

Safeguards often do not deliver the expected result. In 1982, Wilde proposed one potential reason for this in the risk homeostasis theory. This theory holds that whatever a person does, each individual holds a specific level of risk that he or she finds subjectively acceptable. While engaged in a given task, individuals monitor the level of risk they feel exposed to and regulate their behavior so as to keep it at an acceptable level. If the perceived level of risk is lowered through adding a safety mechanism, more risky behaviors will be displayed until the targeted risk level is reached (provided that the risky behavior is attractive). Originally developed in the context of traffic accidents, the theory tried to account for risky driving. In the framework of this theory, driving at a specific speed is associated with a specific probability (i.e., risk) of injury in case of an accident. Reducing this risk level through fitting the car with air bags will increase the speed of driving. In an industrial context, this theory explains why fitting a safety mechanism (e.g., installing a safety screen in front of a galvanic bath) possibly increases hazardous behavior (e.g., being less careful to avoid splashes, not wearing protective goggles). Finally, the third level of prevention is used when the means of design and safeguards are exhausted. Then, warning systems such as signs and symbols are installed.

As pointed out previously, whenever systems use the second and third levels of prevention, safety depends on the extent to which workers work with or alongside the safeguards and safety procedures. This understanding broadens the focus of safety measures from a pure engineering perspective to look at worker behaviors. The majority of behavior-oriented research has focused on compliance to safety rules such as whether workers use the proper work strategies, and use tools properly. In 2002, Burke and colleagues developed a broader taxonomy of generic safety performance with the goal of capturing categories of safety behaviors that cut across different jobs. This research, carried out in the nuclear waste industry, defined general safety performance as "the actions or behaviors that individuals exhibit in almost all jobs to promote the health and safety of workers, clients, the public, and the environment." Four categories of safety performance were identified: (a) using personal protective equipment to shield individuals from hazards, (b) engaging in work practices to reduce risk (e.g., adhering to safety procedures in terms of risk-reducing work practices), (c) communicating health and safety information (e.g., a hazard, an accident, a related incident) to the appropriate personnel, and (d) exercising employee rights and responsibilities. Although previous research that looked at worker behaviors typically emphasized worker safety compliance (as captured in the first two categories), the third and fourth categories identified in this research clearly went beyond this approach. A fruitful addition to this taxonomy, developed by Griffin and Neal in 2000, is the extent to which workers participate voluntarily in safety activities or attend safety meetings. It is assumed that these behaviors, called safety participation, contribute indirectly to workplace safety by supporting the development of a safe environment.

6. RECENT APPROACHES TO SAFETY: THE ROLE OF MANAGEMENT AND ORGANIZATIONAL CLIMATE

In many industrial contexts, unsafe behaviors are inherently more rewarding than safe behaviors. They frequently allow faster and more comfortable accomplishment of work, and they are often more socially acceptable (e.g., the display of safe behavior may be regarded as "unmanly"). Accidents or injuries occur too rarely to represent a sufficient reinforcement for displaying safe behaviors. So, what induces individuals to act in a safe way despite the attractiveness of unsafe behaviors? Characteristics of individuals (e.g., their

safety knowledge) will enhance these behaviors. Of similar importance is a factor that pertains to the environment in which workers operate, that is, the organization's safety climate.

Organizations develop policies, procedures, and practices concerning occupational safety. Organizational climate has been defined as employees' perception of these policies, procedures, and practices. For example, safety climate captures the level of priority accorded to occupational safety, and this is assumed to influence individuals' safety behavior. It is presumed that safety climate is shaped by supervisors as well as by top management (cf. Fig. 1).

Supervisors, especially line managers, influence an organization's safety climate through several mechanisms. First, line managers serve as role models for their staff. In 2000, Maierhofer and colleagues found line managers' use of personal protection equipment to be related to the frequency with which employees use it. Second, also in 2000, Zohar demonstrated that supervisors' tendency to respond to staff safety violations, to react positively to safe behaviors, and to show that they take safety seriously is important to bringing rules and regulations to life. This supervisor behavior also proved to be related to microaccidents. In 1999, Hofmann and Morgeson revealed that the more comfortable employees felt in raising safety concerns and discussing safety issues with their supervisor, the lower the accident rate. The wider social environment existing with coworkers also has some potential to control unsafe behaviors and its consequences. For example, in 1996, Hofmann and Stetzer found, at the group level, that the more inclined individuals in a group were to approach someone about unsafe behaviors, the lower the accident rate.

The Zeebrugge ferry incident showed how apparently minor examples of management behavior (e.g., refusal to install a warning light on the bridge of the ship) can combine with other policies (e.g., negative reporting only) in the antecedents of an accident. These illustrations are indicative of poor safety climates.

The meaning of supervisor behavior becomes more significant when workers are confronted with high performance pressure. Under time pressure, shortcut work methods are more readily adopted to save some time—often at the expense of safety—and supervisor behavior plays a critical role in the alignment of these competing goals of working safely and reaching performance targets. Zohar's study on microaccidents showed that although there was a general trend for microaccidents to increase with growing time pressure, this effect was reduced when the supervisor reinforced safe behaviors and sanctioned unsafe behaviors.

In addition to direct supervisors, top management influences organizational safety by shaping the safety climate and, in turn, workers' safety behaviors. Top management's commitment to safety issues, the priority top management places on safety that shows through allocation of resources and investments (for tools, health, and safety training), and the support that top management gives to direct supervisors all shape the safety climate. Shannon and colleagues' 1997 overview of studies on injury rates showed that injury rates were consistently lower in organizations where top management played an active role in health and safety issues.

7. SAFETY-ENHANCING INTERVENTIONS

Interventions aim at enhancing safety using one of three pathways: changing characteristics of the workforce (e.g., increasing safety behavior), changing the work itself (e.g., improving the design of tools), or altering elements of the work context (e.g., changing reporting principles).

The interventions employed most frequently seek to influence characteristics of workers, mainly to enhance safety behaviors. In 2003, Burke and Sarpy suggested that safety behavior could be seen as a function of knowledge, skills, and abilities (KSAs) as well as motivation and values (cf. Fig. 1). Safety training, based on diverse training methods such as lectures, demonstrations, and hands-on simulations, have been employed successfully in providing staff with the relevant KSAs and in motivating staff to use the skills. Research on training effectiveness outside the domain of safety has identified numerous factors that influence the effectiveness of training, and many of these can be applied to safety training as well.

Other interventions seek to specifically enhance the motivation to display the trained behavior using behavioral principles. One element used by a variety of intervention strategies is the use of feedback that informs workers about their level of safety behavior or number of accidents. An intervention study in the vehicle maintenance division by Komaki and colleagues in 1980 showed that training was somewhat useful in preventing accidents but that a much stronger effect, in terms of reduced accident rates, was achieved

when the training was accompanied by regular feedback to workers on safety-related behavior. Related interventions have successfully merged training with goal setting or used incentives and competition between units—always combined with feedback.

The second type of intervention, changing work itself, is employed whenever engineering control mechanisms are used. Furthermore, changing the pace of work to make safe behaviors less "costly" for workers may be important. Finally, changing the work context can improve safety. For example, in 2001, Kaminski showed that rewarding principles are related to injuries; that is, the use of performance-related pay, an important and successful principle of performance management, was significantly related to higher levels of injuries.

Finally, a more encompassing intervention method would target higher levels of an organization's hierarchy. Changing line, middle, or top managers' approaches to safety can improve the organization's safety climate, and this in turn can enhance employees' safety behaviors. In 2002, Zohar reported such an intervention. The intervention sought to increase the relative frequency with which line managers spoke to their staff about safety behaviors versus production-oriented issues. Increasing the proportion of safety-oriented interaction with a feedback method made safety an explicit performance goal. As a result, employees' safety behaviors improved, their perception of the safety climate increased, and the injury rate was reduced successfully.

8. SUMMARY

This article has traced accidents as having antecedents in the twin areas of error and safety. By understanding these, it is possible to trace causes of accidents and take judicious steps to avoid their recurrence. Combining the engineering approach with the various strategies to increase the display of safety behaviors would seem to be fruitful.

See Also the Following Articles

Environmental Versus Individual Risk Taking: Perception, Decision, Behavior ■ Work Safety

Further Reading

Barling, J., & Frone, M. (2004). *The psychology of workplace safety*. Washington, DC: American Psychological Association.

Burke, M. J., & Sarpy, S. A. (2003). Improving worker safety and health through interventions. In D. A. Hofmann, & L. E. Tetrick (Eds.), *Health and safety in organizations* (pp. 56–90). San Francisco: Jossey–Bass.

Burke, M. J., Sarpy, S. A., Tesluk, P. E., & Smith-Crowe, K. (2002). General safety performance: A test of a grounded theoretical model. *Personnel Psychology*, *55*, 429–457.

Frese, M., & Zapf, D. (1994). Action as the core of work psychology: A German approach. In H. C. Triandis, M. D. Dunnette, & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (2nd ed., Vol. 4, pp. 271–340). Palo Alto, CA: Consulting Psychologists Press.

Griffin, M. A., & Neal, A. (2000). Perceptions of safety at work: A framework for linking safety climate to safety performance, knowledge, and motivation. *Journal of Occupational Health Psychology*, *5*, 347–358.

Hofmann, D. A., & Morgeson, F. P. (1999). Safety-related behavior as a social exchange: The role of perceived organizational support and leader–member exchange. *Journal of Applied Psychology*, *84*, 286–296.

Hofmann, D. A., & Stetzer, A. (1996). A cross-level investigation of factors influencing unsafe behaviors and accidents. *Personnel Psychology*, *49*, 307–339.

Kaminski, M. (2001). Unintended consequences: Organizational practices and their impact on workplace safety and productivity. *Journal of Occupational Health Psychology*, *6*, 127–138.

Komaki, J. L. (1998). *Leadership from an operant perspective*. New York: Routledge.

Komaki, J., Heinzmann, A. T., & Lawson, L. (1980). Effect of training and feedback: Component analysis of a behavioral safety program. *Journal of Applied Psychology*, *65*, 261–270.

Laflamme, L., & Menckel, E. (1995). Aging and occupational accidents: A review of the literature of the last three decades. *Safety Science*, *21*, 145–161.

Ludwig, T. D., & Geller, E. S. (1997). Assigned versus participative goal setting and response generalization: Managing injury control among professional pizza deliverers. *Journal of Applied Psychology*, *82*, 253–261.

Maierhofer, N. I., Griffin, M. A., & Sheehan, M. (2000). Linking manager values and behavior with employee values and behavior: A study of values and safety in the hairdressing industry. *Journal of Occupational Health Psychology*, *5*, 417–427.

Norman, D. A. (1981). Categorization of action slips. *Psychological Review*, *88*, 1–55.

Rasmussen, J. (1983). Skills, rules, and knowledge; signals, signs, and symbols; and other distinctions in human performance models. *IEEE Transactions on Systems, Man, and Cybernetics*, *13*, 257–266.

Reason, J. (1990). *Human error*. Cambridge, UK: Cambridge University Press.

Reason, J. A. (1997). *Managing the risks of organizational accidents*. Aldershot, UK: Ashgate.

Reber, R. A., Wallin, J. A., & Chhokar, J. S. (1990). Improving safety performance with goal setting and feedback. *Human Performance*, *3*, 51–61.

Salgado, J. (2002). The Big Five personality dimensions and counterproductive behaviors. *International Journal of Selection and Assessment*, *10*, 117–125.

Salvendy, G. (1997). *Handbook of human factors and ergonomics*. New York: John Wiley.

Sanders, M. S., & McCormick, E. J. (1993). *Human factors in engineering and design* (7th ed.). New York: McGraw–Hill.

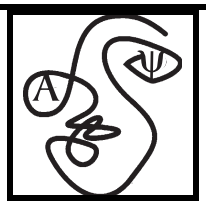
Shannon, H. S., Mayr, J., & Haines, T. (1997). Overview of the relationship between organizational and workplace factors and injury rates. *Safety Science*, *26*, 201–217.

Wilde, G. J. S. (1982). The theory of risk homeostasis: Implications for safety and health. *Risk Analysis*, *2*, 209–258.

Zapf, D., Brodbeck, F. C., Frese, M., Peters, H., & Pruemper, J. (1992). Errors in working with office computers: A first validation of a taxonomy for observed errors in a field setting. *International Journal of Human–Computer Interaction*, *4*, 311–339.

Zohar, D. (2000). A group-level model of safety climate: Testing the effect of group climate on microaccidents in manufacturing jobs. *Journal of Applied Psychology*, *85*, 587–596.

Zohar, D. (2002). Modifying supervisory practices to improve subunit safety: A leadership-based intervention model. *Journal of Applied Psychology*, *87*, 156–163.



School–Community Partnerships

Richard J. Cowan
Kent State University, Kent, Ohio, USA

Susan M. Swearer
University of Nebraska, Lincoln, Nebraska, USA

1. School–Community Partnerships Defined
2. Domains of School–Community Partnerships
3. Approaches to Fostering School–Community Partnerships
4. Conclusion
Further Reading

GLOSSARY

collaboration Working together toward a common goal or set of goals.

community-based organization Any agency, organization, or system (other than a school) that exists in the community and has the potential to partner with a school (e.g., public library, local law enforcement agency, community mental health clinic, local family medical clinic) to meet the needs of children, families, and the community as a whole.

educators General education instructors, special education instructors, teaching aides, classroom volunteers, administrators, support staff, school psychologists, and any other persons within a school setting that contributes to the education of students within that environment (e.g., lunchroom aides).

parent The adult who is responsible for the care and well-being of a child (e.g., biological parent, caregiver, legal guardian).

partnership A coequal interdependent relationship that is established and developed over time with the primary purpose of working together toward a mutually determined set of goals and objectives.

problem solving Process of identifying a priority need or problem, determining outcome goals and objectives, assessing what might be causing and/or maintaining the problem (or blocking access to a desired resource), determining a solution to target potential causal mechanisms, implementing the intervention or strategy to meet the need or solve the problem, evaluating outcomes with regard to predetermined goals and objectives, and determining the need for continued intervention; this process may be recycled as needed based on the assessment of coparticipants.

system A particular environment with an emphasis on specific components of the environment, including people, physical objects, placement of people and objects, physical arrangement, and interactions among these components (e.g., home system, school system).

Partnership may be defined as a coequal interdependent relationship that is established and developed over time with the primary purpose of working together toward a mutually determined set of goals and objectives. In the case of school–community partnerships, the targeted goals and objectives are related to the needs of students and the community as a whole. Such partnerships use individuals and resources from the school, together with those from the community, to serve both the educational system and the community. School–community partnerships are capable of addressing multiple domains that include, but are not limited to, family and student health, social and mental health programs, violence prevention, tolerance and diversity, government legislation/regulation and law enforcement, community volunteers and the faith community, business–community partnerships, and school/academic academic readiness and educational support.

©2004 Elsevier Inc.
All rights reserved.

1. SCHOOL–COMMUNITY PARTNERSHIPS DEFINED

1.1. Partnerships Defined

A common definition of the term "partnership" is that of a coequal interdependent relationship that is established and developed over time with the primary purpose of working together toward a mutually determined set of goals and objectives. As applied to relationships between schools and communities, a partnership generally refers to a relationship occurring between educators (e.g., teachers, school psychologists, support staff) and parties from noneducational settings (e.g., parent and advocacy groups, mental health care workers from community wellness centers, doctors from local clinics and hospitals, law enforcement officers), with an emphasis on goals and objectives related to the educational and social development and the well-being of students for whom all members of a particular community share interest and responsibility. In addition, school–community partnerships are formed to address the needs of the community as a whole. One could easily argue that because students ultimately grow up and become members of the community, school–community partnerships that target the needs of students within the school serve the interests of both developing students and the future of the community. Establishing and developing a partnership in which participants share ideas, information, resources, skills, and expertise is a process that requires care and cultivation. Partnerships do not result from one or two encounters between educators and noneducators within the community; rather, they develop over time through multiple collaborative interactions. The process of building and developing partnerships requires a guiding philosophy. Approaches to fostering school–community partnerships are discussed later in this article, following a discussion of characteristics of, parties served by, and domains for activities related to school–community partnerships.

1.2. Characteristics of School–Community Partnerships

In 2002, Sheridan and colleagues offered the following characteristics of meaningful school–community partnerships:

1. Interactions among partners are collaborative.
2. Maintenance of a positive relationship is a priority.
3. Relationships are balanced.
4. Relationships are cooperative and interdependent.
5. Relationships occur within a context with students at the center.
6. Actions between schools and other community agencies are flexible, responsive, and proactive.
7. Differences in viewpoints and perspectives are seen as strengths rather than hindrances.
8. There is a commitment to cultural diversity.
9. There is an emphasis on outcomes and goal attainment.

The first characteristic, collaborative interactions, may take place through collaborative problem solving whereby there is mutual determination of needs and goals in addition to joint ownership of solutions and outcomes. Positive relationships may be defined as relationships that are balanced, cooperative, interdependent, and equal in terms of decision-making power. Because schools and other community agencies serve multiple unique functions in a given community, it is important for partners to remain flexible, responsive, and proactive. It is nearly certain that differences in opinion regarding priorities and strategies will result during interactions among various individuals representing different agencies. However, if the partners embrace a coequal partnership approach, whereby each participant's relative expertise is valued and equal decision-making power is granted, these differences are not viewed as hindrances; rather, they are viewed as opportunities to work through differences in a joint effort to meet the needs of students and families served by the community. School–community partners should consider the needs of all community members, including those individuals and families who come from culturally diverse backgrounds. Cultural diversity extends beyond race and ethnicity to include family attitudes, beliefs, values, and experiences with various schools and community-based organizations. Shared ownership and responsibility for the education, development, and well-being of all children allows for the whole to be greater than the sum of its parts and is derived from the work of individuals who are able to put aside their personal needs and goals in an effort to meet the needs and goals of the entire community.

1.3. Foci of School–Community Partnerships

In 2002, Epstein and colleagues, who are leaders in the area of family–school–community partnerships within the field of school psychology, maintained that school–

community partnerships generally center on the needs of four identifiable targets: (a) student-centered services, (b) family-centered services, (c) school-centered services, and (d) community-centered services. Student-centered partnership activities may focus on the development of student awards, student incentives, scholarships, student-centered extracurricular activities (including travel), and scholastic and career mentoring. Family-centered partnership activities may focus on parent training and education, family-based recreational activities, adult educational courses (e.g., GED courses), parent rewards, parent counseling and consultation, and other contributions for parents. School-centered partnership activities may focus on gaining access to equipment and materials, beautification and repair of the school and/or community, teacher-related rewards and incentive programs, funds for school activities and events, and office and classroom supports. Finally, community-centered activities implemented through school–community partnerships may focus on community improvement and beautification, student-generated community-supported exhibits and performances, volunteer work, and other outreach activities.

2. DOMAINS OF SCHOOL–COMMUNITY PARTNERSHIPS

2.1. Ecological Theory

The fields of education and psychology have come to embrace a systems-based approach to understanding those variables that influence how an individual behaves, develops, and learns over time and across multiple settings. This approach is grounded in ecological theory. Based on the work of Bronfrenbrenner, ecological theory is concerned with the interaction between an individual and various contextual systems. Systems include the microsystem, mesosystem, exosystem, and macrosystem. The microsystem is concerned with the individual and his or her immediate environment (e.g., a child interacting within a particular classroom), the mesosystem involves the interrelation between major systems in the individual's life (e.g., between home and school), the exosystem is concerned with environments not directly related to the individual yet influencing his or her life (e.g., a parent's workplace), and the macrosystem includes overall cultural or subcultural patterns and influences (e.g., policies, federal and state legislation, national and global economic factors). This ecobehavioral paradigm illustrates how systems influence and are influenced by one another, stressing that contextual systems do not exist in isolation; rather, interactions across systems occur frequently and are multifaceted and multidetermined over time. All system levels influence school–community partnerships, with multiple mesosystemic intersections occurring between the school and various community settings. For example, partnerships may exist between a school and a business within the community or between the school and a community mental health agency. These intersections between the school and various community settings involve multiple individuals, are influenced by multiple factors, and represent opportunities for both educational and noneducational partners to intervene with regard to multiple domains. Some of these domains are discussed next.

2.2. Specific Domains

Specific domains to consider for potential partnership activities include family and student health, social and mental health programs, violence prevention, tolerance and diversity, government legislation/regulation and law enforcement, community volunteers and the faith community, business–community partnerships, and school/academic readiness and educational support (Table I). These domains of school–community partnerships are not mutually exclusive.

2.2.1. Family and Student Health

Because good health has implications for education and development, and because learning about wellness and a healthy lifestyle is a critical curriculum in and of itself, schools and other community-based

TABLE I
Domains of School–Community Partnerships

- Family and student health
- Social and mental health programs
- Violence prevention
- Tolerance and diversity
- Government legislation/regulation, and law enforcement
- Community volunteers and the faith community
- Business–community partnerships
- School/academic readiness and educational support

organizations (e.g., health care clinics) are identifying opportunities for partnerships with the primary goal of wellness education (often as a preventive mechanism) in addition to medical intervention for prevention (e.g., flu shots, immunizations) and remediation (e.g., medical intervention) of medical problems. Bringing wellness education and medical services to the school makes sense given that making appointments is a challenge for many families and given that the school setting is where most children spend the majority of their waking hours. School-based nurses, counselors, and/or psychologists are in an ideal position to collaborate with medical health care partners to ensure treatment fidelity and monitor behaviors related to medical intervention that may be indicators of effectiveness (e.g., monitoring the effects of medication on the on-task performance of an individual diagnosed with attention deficit/hyperactivity disorder).

2.2.2. Social and Mental Health Programs

The field of school psychology has experienced an enormous shift in the provision of services, expanding beyond the role of conducting psychoeducational assessment to include roles such as consulting with parents and educators, being involved in program development and evaluation, providing small group interventions for students with special needs (e.g., social skills training), crisis intervention, and working with community-based organizations to form partnerships in the best interest of individuals and families served by the school and community. School psychologists strive to expand their role as practitioners in schools and communities to provide a wider range of mental health services to individuals and families within the school. This could include direct services (e.g., therapy, one-on-one counseling), indirect services (e.g., behavioral consultation with parents and educators, family-based consultation services), and group-based services (e.g., anger management groups, social skills training groups, mental wellness training programs).

School psychologists, school counselors, and social workers may be in an ideal position to develop, implement, and evaluate mental health services. As with the provision of medical services within the school setting, providing mental health services within the school has implications for the education and development of students and makes common sense given that school-aged children spend a significant proportion of their waking hours within the school setting. Providing a wider array of mental health services within the school may require additional training of some school-based personnel. This may be promoted in part through partnerships with professionals from the community (e.g., physicians, licensed psychologists, mental health care providers) whereby experts provide training and professional development in specific domains of mental health services, including wellness training and preventive mental health care practices.

2.2.3. Violence Prevention

Given the recent national focus on violence prevention and intervention among students, schools have been encouraged, or even mandated, to implement programs that tackle the problem of violence in schools. Violence occurs on a continuum ranging from bullying to more extreme behaviors such as assault and homicide. Although most schools have adopted a "zero tolerance" policy toward more severe forms of violence, schools are also extending this hardline approach to bullying. In fact, as Limber and Small noted in 2003, there are currently 15 states that have passed laws to address bullying in schools. These laws have set the stage for parents to pursue litigation if they believe that their children are being bullied at school.

School–community partnerships are vital for reducing violence among and toward students. In fact, many schools have law enforcement officers who help to ensure the safety of students in and around the school buildings. Partnering with local law enforcement is an important component in helping to reduce violence in schools. In addition, it is clear from the research that violence prevention is a whole school and community responsibility. Violence prevention programs must have buy-in from students, teachers, administrators, parents, and adults in the community. When school–community partnerships are forged, violence prevention efforts can be successful.

2.2.4. Tolerance and Diversity

The educational system in the United States continues to experience an increase in culturally diverse students and families. Some schools that have experienced an increase in culturally diverse students and families have found it helpful to seek partnerships with community-based organizations to help families acclimate to the school system and community that they are entering.

For example, there are community-based organizations whose primary mission is to provide English as a Second Language (ESL) or English as a Learned Language (ELL) programs for students and families as well as to provide transition supports to families and schools as they learn to mutually accommodate one another's attitudes, beliefs, and values.

In addition to seeking partnerships with community-based organizations that help families in transition become a part of the educational and community systems, schools may seek interactions with organizations that serve particular minority groups in an effort to expose students to new cultures and diversity. This latter type of partnership affords school-based learning opportunities for students to interface with individuals with cultural traditions that are different from their own in hopes of instilling greater acceptance and tolerance. Such partnerships may also allow educators to gain insight into a family's attitudes, beliefs, and values regarding education and citizenship.

2.2.5. Government Legislation/Regulation and Law Enforcement

The U.S. educational system is replete with state and federal legislation that guides the provision and implementation of educational services and programming. One implication of such legislation is that the school is responsible for providing a free and appropriate public education (FAPE) to all individuals, including those with specific learning, medical, and/or psychological disabilities. Once an individual is identified with a specific disability, the educational team must determine specific learning goals and objectives related to the needs of that individual. Once these goals and objectives are determined and specific accommodations are specified, the school must provide a meaningful placement for the student to learn. The school is legally required to provide services in the least restrictive environment (LRE) possible, meaning that whenever possible and reasonable, the student should be included in the general education setting with peer models to support his or her learning. In addition, once a student meets eligibility for special education services, the school is responsible for providing services until the student is 21 years of age. Some high school-aged students will require accommodations and placements that foster the development of independent living and vocational skills. Because school settings might be limited in terms of vocational training opportunities, schools often seek partnerships with community-based organizations to place students in industry-specific training programs. Ideally, the school determines each student's relative strengths and needs, conducts a thorough assessment of abilities and interests, and seeks to match these variables with specific job training sites. This type of partnership not only addresses the needs of individual students and families but also helps contribute to future workers who are needed in community-based employment programs.

There is a long-standing tradition of partnerships between schools and community-based law enforcement agencies whereby police officers come into schools to talk about crime and drug abuse prevention. In many cases, local law enforcement agencies provide curricular materials to supplement their training programs beyond 1- or 2-day presentations. These types of programs generally involve the use of multimedia presentations and/or live skits to demonstrate important safety-related tips. At the middle school and high school levels, it is common to have law enforcement officers placed within a school building to ensure that students abide by the law as well as to provide opportunities for police officers to teach classes geared toward recruiting high school seniors into law enforcement-related occupations.

2.2.6. Community Volunteers and the Faith Community

Because most communities have unmet needs (e.g., streets and parks that need to be cleaned, elderly individuals who might not have many social opportunities, citizens who cannot afford clothing and/or food), and because many view the educational process as an opportunity to instill ethical and character-building beliefs and practices, fostering volunteering opportunities for students within local faith-based and/or charitable organizations represents a unique opportunity for students to learn the value of making contributions to the community while at the same time seeing that a need of the community is met, at least for a brief time. This type of partnership will vary greatly across schools and regions of the country, being heavily dependent on the predominant faith of a particular community and/or whether the school is private or public. However, a common theme of school–volunteer agency partnerships is that students make a contribution to a local organization that is based on volunteering and providing services to the community while at the same time benefiting from such an experience.

2.2.7. Business–Community Partnerships

Because major companies need tax breaks, and because making contributions to schools allows for an investment in the future of the local community, school–business partnerships have the potential to be a "win–win" situation. The range of activities may vary greatly; however, there is a common theme whereby businesses "adopt-a-school" as an "educational partner." At this point, the business may make financial contributions, offer training opportunities and/or paid jobs to select students, contribute materials and equipment that may be business related (e.g., computers, calculators), and/or offer to pay employees to volunteer within the school setting. This type of partnership provides materials, training, and support to the school while at the same time building a potential pool of future employees and/or customers for the business partner.

2.2.8. School/Academic Readiness and Educational Support

An important aspect of academic readiness is the ability to read. In addition to partnering with parents to foster regular reading in the home learning environment, schools frequently seek support from local community agencies such as libraries and teacher training programs within local colleges and universities with an emphasis on securing materials and resources to help provide supplemental reading programs. There are many opportunities for schools to partner with other community-based organizations to help enhance and promote primary curricular skills and activities outside of the classroom. For example, schools may partner with local city and county library systems to encourage and foster summer reading programs, including incentives for students and families who participate in active reading throughout the out-of-school months. Another means of enhancing reading skills in students is to seek opportunities to include teachers-in-training from local and regional colleges and universities in reading enhancement programs. This form of partnership allows teachers-in-training to learn about classroom management and literacy curricula while at the same time providing an invaluable service to students in the community. Reading is not the only subject area that may be targeted in school–community partnerships emphasizing academic skills support. For example, schools may partner with local natural history museums and/or planetariums to provide science-based extracurricular activities with incentives for child and family participation outside of the school setting.

3. APPROACHES TO FOSTERING SCHOOL–COMMUNITY PARTNERSHIPS

3.1. Guiding Principles

In 2001, Christenson and Sheridan maintained that partnerships are derived from the careful consideration of four primary components: approach, attitudes, atmosphere, and actions. Although this particular partnership model was developed to describe home–school partnerships, the core components of these authors' philosophy are easily translated into a philosophy that may guide school–community partnerships. Approach is a means of interacting with another party that embraces the value of both parties with regard to the educational and social development and well-being of students for whom all participants share interest and responsibility. This component is concerned with communicating to the other party that its members possess unique knowledge and skills that make them invaluable partners. Attitudes are the perceptions that participants in partnerships have of one another. This is concerned with identifying strengths in partners, focusing on these relative strengths, and assuming a joint responsibility for the education of students within a community. Atmosphere is the overall climate set for participants engaged in a partnership. Finally, actions are strategies—not activities—intended to build long-term relationships through actively involving all partners in the educational process. This philosophy offers specific components for partnership facilitators to consider when working with individuals across various systems who possess the intent of establishing and developing long-term collaborative partnerships.

3.2. Roles and Responsibilities

Partnerships may include multiple individuals, including parents, educators (i.e., general education instructors, special education instructors, teaching aides, classroom volunteers, administrators, school psychologists, and other support staff), paraprofessionals, community partners (e.g., medical professionals, mental health professionals, business owners), and students (especially high school students). When partnerships are grounded in relationships that are positive, balanced, cooperative, and interdependent, and that maintain equal power in decision making, participants recognize, appreciate, and welcome the unique relative expertise and skills

of each participant. Each individual's contribution to the partnership-based change process will vary depending on his or her unique set of strengths, skills, and expertise. For example, parents and other family members offer unique information about the home system from which each student is derived. This is important in understanding the relative contribution of each setting to the behavior and development of students. Parents may intervene to assimilate community values in the home. In addition to providing information about the school setting, educators offer much knowledge in terms of general learning principles, specific academic areas, and various curricula. Their participation may be critical to determine means of implementing interventions within the educational system. Depending on their background, paraprofessionals and community partners offer unique expertise in specific domains that fall outside of parental roles and educational programming per se. Their involvement can provide critical connections to various community-based systems and organizations. In addition, they may take on roles directly related to their practice (e.g., medical doctors offering health and wellness checkups, mental health professionals offering student and/or family therapy). Finally, older students make invaluable partners in determining a valid needs assessment and in determining specific means of proactive student involvement and/or response to adult-generated priorities and strategies. If long-lasting systems level change is to occur, it is critical to involve key players from multiple interconnected systems to assist with implementing specific strategies that will alter the attitudes, values, and beliefs of multiple community members.

3.3. Strategies for Change

Establishing and developing school–community partnerships between one or more parties requires that an individual or a small team takes the initiative in organizing and promoting cross-setting, multipurpose relationships between schools and other community-based organizations. Such planned facilitation allows for the process to unfold in a predetermined organized manner. Those individuals who guide the organizational process must take precautions against being too directive, thereby resulting in an imbalance of power across partners. The initial step in the process is to develop a working relationship between educators and community-based noneducators. During this phase, it is critical to strive for relationships that are positive, balanced, interdependent, and based on coequal status regarding decision-making power. Seeking partners in the community is an ideal role for a school psychologist or administrator who possesses multiple connections in the community and with families served by the school. Initiating a long-term collaborative relationship requires a guiding philosophy. The "four A's" (i.e., approach, attitudes, atmosphere, and actions) derived from Christenson and Sheridan's model is one such philosophy. This includes recruiting various individuals from community-based settings to participate in an inviting environment. If relative expertise is recognized and valued, each member will make a unique contribution to the process. The contribution of the facilitator is expertise regarding how to establish and develop collaborative interactions across individuals and settings (i.e., knowing how to establish partnerships). This person can explain the process, clarify roles, and provide a collaborative open means of communicating and interacting with all participants. Once relationships are formed and the process is clarified and agreed on by all participants, the team may move forward with the team-based problem-solving process.

The team-based problem-solving process generally occurs through six predetermined phases: (a) problem identification, including needs assessment; (b) determination of outcome goals and objectives; (c) ecological assessment, including identification of resources and limitations; (d) intervention development; (e) intervention implementation; and (f) evaluation of outcomes (Table II). Although these phases are sequential, the process is dynamic and may be recycled as needed to meet the mutually determined goals and objectives of the project. The first step in this problem-solving process is to identify a priority need or problem. This step is commonly referred to as a school–community needs assessment. The team should include as many relevant participants from across settings as possible so as to increase the likelihood that multiple individuals and settings will commit to change. A needs assessment may occur through interviews, surveys, and/or group meetings. This step requires that participants put aside their own desires and needs in an effort to focus on specific needs or problems that affect the school and community as a whole. Once a need or problem is targeted, the team must mutually determine specific, observable, measurable outcome goals and objectives. Because progress toward meeting these goals and objectives will serve as a means of assessing overall success, the team must strive to be specific and reasonable. Identifying specific goals allows the team to determine how to measure progress toward goal attainment using

TABLE II
Phases of the School–Community Partnership Problem-Solving Process

1. Problem identification, including needs assessment
2. Determination of outcome goals and objectives
3. Ecological assessment, including identification of resources and limitations
4. Intervention development
5. Intervention implementation
6. Evaluation of outcomes

objective concrete units of analysis. Identifying reasonable goals helps to foster initial success and often involves first focusing on short-term, readily attainable goals and then expanding toward long-term, "broader picture" goals.

The next step is to conduct an ecological assessment to determine what might be causing and/or maintaining a problem or blocking access to a desired resource. This assessment involves examining the characteristics of the school, community, individuals, and families to determine resources they may contribute to an intervention. Resources might include individual expertise, financial contributions, materials, access to research and training, and/or existing interventions. Once the team has (a) made hypotheses regarding what might be causing or maintaining a problem and (b) identified resources available in the school and community, the team should mutually determine the type of intervention to be implemented. As the intervention is developed, it is critical to clarify the intervention across all participants and settings and to assign specific roles to ensure that the plan is carried out as intended. In addition to clarifying roles, it is helpful to write out the plan and make sure that everyone has a copy of the written product. It is often beneficial to assign a person(s) to monitor treatment fidelity throughout the intervention. This person can ensure that the treatment is being implemented as intended as well as offer support as needed. The final step in the process is to evaluate outcomes with regard to predetermined goals and objectives and to determine the need for continued intervention. At this point, the team evaluates the need or problem with regard to change over time, with the goals and objectives serving as "bars" for measuring relative change and/or overall success. If the initial goal is met, the team may decide how to adjust the existing intervention such that a larger goal may be met. If the initial goal is not met, the team may decide to change the intervention and/or recycle through the problem-solving sequence.

4. CONCLUSION

By branching out and fostering long-term relationships with pivotal community partners, schools can create a team oriented toward student-, family-, school-, and/or community-centered services. Establishing and developing a partnership in which participants share ideas, information, resources, skills, and expertise is a process that requires care and cultivation. Such relationships are attainable when all partners value equal decision making, approach each other with mutual respect, work toward the development of attitudes that highlight the strengths of all participants, develop a collaborative inviting atmosphere, and engage in actions that promote meaningful involvement of all participants. Partnerships do not result from one or two encounters between educators and noneducators within the community; rather, they develop over time through multiple collaborative interactions (e.g., the problem-solving process). To meet long-term, mutually determined goals and objectives, it might be necessary for the team to recycle through this change process until all parties are satisfied that the need has been met or the problem has been solved with regard to the mutually predetermined set of goals and objectives. Long-term partners are likely to engage in this problem-solving process multiple times to meet the needs of the school and community. School–community partnerships are critical given the multiple and ever-evolving needs of the schools and communities in which individuals function on a regular basis.

See Also the Following Articles

Community Psychology ■ Consultation Processes in Schools ■ Family and Culture ■ Home–School Collaboration ■ School Violence Prevention

Further Reading

Bronfrenbrenner, U. (1979). *The ecology of human development*. Cambridge, MA: Harvard University Press.

Christenson, S. L., & Sheridan, S. M. (2001). *Schools and families: Creating essential connections for learning*. New York: Guilford.

Epstein, J. L., Sanders, M. G., Simon, B. S., Salinas, K. C., Jansorn, N. R., & Van Voorhis, F. L. (2002). *School,*

family, and community partnerships: Your handbook for action (2nd ed.). Thousand Oaks, CA: Corwin.

Heneggler, S. W., Schoenwald, S. K., Borduin, C. M., Rowland, M. D., & Cunningham, P. B. (1998). *Multisystemic treatment of antisocial behavior in children and adolescents*. New York: Guilford.

Howard, K. A., Flora, J., & Griffin, M. (1999). Violence prevention programs in schools: State of the science and implications for future research. *Applied and Preventive Psychology*, *8*, 197–215.

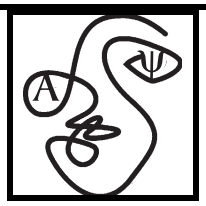
Larson, J., Smith, D. C., & Furlong, M. J. (2002). Best practices in school violence prevention. In A. Thomas, & J. Grimes (Eds.), *Best practices in school psychology* (4th ed., pp. 1081–1097). Bethesda, MD: National Association of School Psychologists.

Limber, S. P., & Small, M. A. (2003). State laws and policies to address bullying in schools. *School Psychology Review*, *32*, 445–455.

Nelson, G., Amio, J. L., Prilleltensky, I., & Nickels, P. (2000). Partnerships for implementing school and community prevention programs. *Journal of Educational and Psychological Consultation*, *11*, 121–145.

Sheridan, S. M., Napolitano, S. A., & Swearer, S. M. (2002). Best practices in school–community partnerships. In A. Thomas, & J. Grimes (Eds.), *Best practices in school psychology* (4th ed., pp. 322–336). Bethesda, MD: National Association of School Psychologists.

Welch, M., & Sheridan, S. M. (1995). *Educational partnerships: Serving students at risk*. Fort Worth, TX: Harcourt Brace.



School Discipline and Behavior Management

George Bear

University of Delaware, Newark, Delaware, USA

1. Introduction
2. Definitions and Components of School Discipline
3. Developing Self-Discipline
4. Using Discipline to Correct Misbehavior
5. Preventing Misbehavior with Effective Classroom Management
6. Remediating Serious Misbehavior
Further Reading

GLOSSARY

alternative education setting A setting in which students are placed as a consequence of serious or chronic misbehavior. The setting is an alternative to the regular classroom and to expulsion.

authoritative discipline A style of discipline that is neither overly harsh nor overly permissive. Authoritative teachers establish clear rules and enforce them in a fair and consistent fashion, hold high behavioral expectations, demonstrate warmth and support, and foster self-discipline.

character education Strategies and techniques that develop thinking, emotions, and behaviors associated with important values and virtues, including self-discipline.

classroom management Practices, procedures, and routines used to establish and maintain a learning environment that fosters instruction and learning.

self-discipline Internal regulation of one's behavior. It entails assuming responsibility for one's actions, understanding right from wrong, and inhibiting socially inappropriate behavior.

zero tolerance approach An authoritarian approach to school discipline in which students are punished harshly, automatically, and irrespective of the circumstances involved.

School discipline and classroom management have always been primary concerns among educators. This article addresses the four major components of effective school discipline and classroom management: developing self-discipline, correcting misbehavior, preventing misbehavior with effective classroom management, and remediating serious and chronic behavior problems.

1. INTRODUCTION

Since the onset of schooling, educators have been concerned about issues of school discipline and classroom management. Educators in Colonial America faced problems of fighting, noncompliance, truancy, and dropping-out—problems common in schools today. To be sure, the frequency and severity of discipline problems have changed over the years. In general, discipline problems have become more common and violent in nature, particularly since the late 1960s. Indeed, nearly every year since an annual Phi

©2004 Elsevier Inc.
All rights reserved.

Delta Kappa/Gallup Poll of the public's attitudes toward public schools was first conducted in the 1970s, the poll has revealed that the "lack of discipline/more control" is either the greatest problem or among the greatest problems facing U.S. schools.

2. DEFINITIONS AND COMPONENTS OF SCHOOL DISCIPLINE

Among the many definitions of discipline, two are most applicable to the context of schooling: (i) teaching that develops self-control and character and (ii) treatment that corrects or punishes (*Webster's New World College Dictionary*, 1996, p. 391). Both of these definitions reflect two traditional goals of U.S. educators: (i) to teach, or develop student self-control, character, or self-discipline, and (ii) to use discipline to correct, punish, or manage misbehavior.

Closely related to school discipline is classroom management. Classroom management refers to practices, procedures, and routines used to establish and maintain a learning environment that fosters instruction and learning. This consists of such tasks as organizing the classroom and materials; managing academic assignments; planning and delivering effective instruction; monitoring student behavior; and establishing rules, procedures, and routines. As is true with the development of self-discipline, classroom management serves a preventive function. That is, both help prevent many discipline problems from occurring in the future. They differ in focus, however. Whereas developing self-discipline focuses primarily on the competencies of students, preventing discipline problems with effective classroom management focuses primarily on the competencies of teachers. Developing self-discipline, using discipline to correct misbehavior, and preventing discipline problems with effective classroom management are important components of a comprehensive school discipline plan. Each of these components should target all students. A fourth component of comprehensive school discipline, remediating serious misbehavior, primarily concerns schoolwide discipline and focuses on the needs of the small percentage of students who typically require removal from the regular classroom and placement in another educational setting. These four components are the focus of this article.

3. DEVELOPING SELF-DISCIPLINE

Until relatively recently, developing self-discipline was the primary aim of U.S. education. The aim was to develop self-governing, socially responsible citizens who require the minimum of external control. Self-governance, or self-discipline, was viewed as necessary for the survival of a democratic society. For several centuries, religious education and moral education served this aim. These two approaches to developing self-discipline were replaced by character education in the early 20th century. Although their methods differed, each of these educational approaches was characterized by systematic and deliberate attempts to teach self-discipline. In nearly all schools, students were required to read value-laden literature, memorize scripts or codes that endorsed religious and moral principles or character traits, discuss issues of morality and problems of democracy, and participate in activities that promoted cooperation, helping, and democratic thinking.

In the 1960s and 1970s, character education faded precipitously. Although student protests and violence in society increased markedly during this period, few schools continued to take deliberate and systematic actions to develop self-discipline. Instead of attempting to develop character, many schools adopted the perspective of cultural relativism, believing that no set of values should be taught in the schools. Instead, it was widely held that students were to develop their own values, with little or no direction from teachers. This perspective, combined with increased concerns about the legal rights of students and parents (especially freedom of speech and the right to due process) and a rapidly growing focus on the teaching of math and science, contributed to the decline of character education.

In many schools, renewed efforts are being made to develop self-discipline. These schools recognize that promoting self-discipline, or responsible citizenship, reduces discipline problems in the classroom, promotes positive relations with others, fosters academic achievement, and enhances self-perceptions of self-worth. These positive outcomes benefit not just individuals and classrooms but also society in general.

Popular programs that promote self-discipline are commonly referred to as character education or social and emotional learning programs. These programs promote self-discipline by targeting a combination of social cognitions, emotions, and social skills that research has shown to be related to prosocial and antisocial behavior. These include empathy, social

perspective taking, attributions of responsibility, social problem solving, interpersonal understanding, moral reasoning, conflict resolution, anger management, and impulse control. In teaching and developing self-discipline, these programs rely heavily on the techniques of verbal instruction (e.g., direct teaching of social problem-solving steps), role playing, modeling, positive practice, positive reinforcement, and classroom discussion of moral and social problems. Such programs teach students how to think, feel, and act—a strategy that most authorities in school discipline believe is much more effective in preventing behavior problems than are strategies designed to teach students what to think and do, such as lecturing, moral appeal, and fear arousal.

Most programs of character education or social and emotional learning are of one of two general types or a combination thereof: a curriculum package or curriculum infusion. A curriculum package, commonly produced by a commercial or nonprofit publishing company, consists of lessons that are typically taught by the classroom teacher at least once a week. Some also include lessons, activities, or support materials for parents. Unfortunately, few of these programs have been evaluated and shown to be effective, especially in reducing or preventing actual misbehavior. There are exceptions, however. Thus, before adopting any program it is strongly recommended that a school review evidence of the program's effectiveness [the Collaborative for Academic, Social, and Emotional Learning has an excellent Web site (www.casel.org) that includes reviews and research on popular social and emotional learning programs].

The second general type of program for developing self-discipline, curriculum infusion, entails integrating lessons and activities for promoting self-discipline throughout the curriculum as well as other school activities. This includes using techniques, such as role playing and class discussion of social problems and moral dilemmas, to foster social and emotional learning when covering many issues in social studies (e.g., constitutional rights, issues of slavery, and the Holocaust), literature (e.g., the moral dilemmas and interpersonal problems of characters in literature), and health (e.g., issues of smoking, alcohol use, and peer pressure). It also includes the use of creative writing, plays, school assemblies, and club activities.

Perhaps the most comprehensive and research-supported program for fostering social and emotional learning is the Child Development Project. This project employs each of the techniques discussed previously while emphasizing the active participation of students in classroom and schoolwide activities that call for social decision making, moral reasoning, and prosocial behavior. The project also emphasizes what it calls developmental discipline as a method of correcting misbehavior and preventing future misbehavior by developing self-discipline. Although developmental discipline includes punitive consequences for misbehavior, it focuses on teaching students to understand and respect the perspectives of others, especially how their own misbehavior impacts the feelings of others. It also focuses on teaching students skills of social problem solving and conflict resolution.

In using developmental discipline, misbehavior is viewed as providing teachers with opportunities to help develop self-discipline. Research shows that schools that participate in the Child Development Project, especially those that employ systematically and consistently a combination of the previously discussed techniques, are effective in promoting social decision making, social problem solving, empathy, respect for the opinions of others, democratic values, and a positive school climate. Moreover, in developing self-discipline these schools reduce future alcohol use and delinquent acts (i.e., threatening others, possessing a weapon, stealing vehicles, and skipping school).

4. USING DISCIPLINE TO CORRECT MISBEHAVIOR

Educators have always used discipline to manage, control, and punish misbehavior. Fortunately, today's methods tend to be less harsh, at least physically, and more sensitive to students' rights than in earlier centuries. Whereas disobedient students were commonly whipped or spanked in the past, approximately 30 states now ban the use of corporal punishment in schools. Unfortunately, many schools have replaced corporal punishment with perhaps an equally harsh "zero tolerance" approach to discipline. A pervasive zero tolerance approach differs from reasonable zero tolerance policies in that the latter applies only to the most serious offenses that threaten the welfare and safety of others. A zero tolerance approach requires that misbehaving students are suspended, expelled, and/or turned over to the police for common crimes and misbehaviors that were typically handled by schools in the past, including fighting, stealing, and vandalism—the three most common criminal offenses in schools. Indeed, the number of suspensions, expulsions, and police reports has increased dramatically

during the past decade, with a greater proportion of African American students being removed from school. Such disproportionality raises important ethical and legal issues. The increase in a zero tolerance approach to school discipline has occurred despite the lack of evidence that this approach, especially when used as the primary means of managing and correcting misbehavior, is effective in developing self-discipline and replacement behaviors or in reducing the future occurrence of the behavior problem. Not only are these practices ineffective but also they often reinforce the very behaviors that led to the student's removal from school. For example, this happens when students who dislike school, teachers, or peers are sent home and, in the absence of parental management and supervision, they spend their time watching television, playing video games, roaming the streets, etc. instead of acquiring academic, vocational, and interpersonal skills.

In responding to more frequent, and less serious, behavior problems, such as a student talking or getting out of his or her seat without permission, educators tend to use punitive, or reductive, techniques that are less harsh than those of suspension and corporal punishment. Most popular among these reductive techniques are redirection (e.g., "John, you should be working on your math now"), ignoring the misbehavior, using a visual or physical cue (e.g., making eye contact, moving near the student, giving a hand signal, and placing a hand on the student's shoulder), verbal warnings and reprimands, taking away privileges, time-out (e.g., requiring that the misbehaving student spend a brief amount of time in a setting away from peers), overcorrection (e.g., requiring a student who writes on his or her desk to wash not only his or her own desk but also those of others), and response cost (e.g., awarding a student 15 minutes of free time but then deducting 1 minute for every incidence of a given inappropriate behavior).

Although easy to use and often effective in stopping misbehavior, at least in the short term, these reductive techniques are subject to the same limitations that characterize all forms of punishment, including corporal punishment, suspension, and expulsion. That is, punishment (i) teaches students what not to do instead of how to behave appropriately; (ii) is nonlasting in effectiveness; (iii) teaches students to punish others; (iv) fails to address factors that underlie or contribute to the punished misbehavior; (v) harms the teacher–student relationship; (vi) is likely to produce undesirable side effects, including resentment, revenge, fear, social avoidance, and poor motivation to learn; and (vii) creates a negative classroom environment.

Despite these limitations, it is perhaps understandable why punishment remains the most commonly used technique for managing and correcting misbehavior: Supported by most parents and schools' code of conduct, nearly all schools require that misbehavior not go unpunished. Moreover, punishment is easy to use. Often, it is effective in immediately stopping misbehavior. Also, it does communicate to all students which behaviors are not acceptable. Although punishment often has these advantages, many authors of popular models of school discipline recommend strongly against the use of punishment. Based more on research, however, other authorities and researchers tend to recommend that mild forms of punishment (not including corporal punishment and school removal) should not be abandoned but be used wisely. That is, it is recommended that punishment be used only after other more positive techniques have already been used and have failed to improve appropriate behavior. Moreover, it is recommended that punishment be used only in combination with corrective techniques designed to teach students to make responsible choices and to exhibit appropriate behavior. Popular corrective techniques that have been shown to be effective in promoting appropriate behavior include individual and class discussions that encourage social problem solving and social decision making, modeling, and the use of various forms of positive reinforcement such as social reinforcement (e.g., teacher praise, peer recognition, and notes sent home to parents), privileges (e.g., recess and parties), and material rewards (stickers, prizes, and certificates) contingent upon good behavior. Not only do these techniques increase appropriate behaviors but also they often simultaneously decrease the frequency of inappropriate behavior. For example, when students are praised or otherwise reinforced for working on their assignments and considering the needs of others, they are less likely to misbehave.

5. PREVENTING MISBEHAVIOR WITH EFFECTIVE CLASSROOM MANAGEMENT

What best differentiates the most effective teachers from the least effective teachers are not the techniques they use to correct behavior problems but, rather, the techniques they employ to prevent behavior problems from occurring in the first place. Effective classroom management not only prevents most behavior problems but also helps

promote the development of self-discipline and academic achievement. It does so by creating a classroom climate characterized by clear procedures, routines, and high expectations, as well as feelings among students of safety, pride, and interest in learning and of warmth, acceptance, and support from the teacher and classmates.

As described in classic studies by Jacob Kounin in the 1960s and 1970s, effective classroom managers demonstrate what Kounin called withitness (they closely monitor the behavior of every student in the class), overlapping (they deal effectively with multiple events simultaneously), smoothness and momentum in lessons (they provide instructions that motivate students), and group alerting (they elicit and maintain the attention of all students). Recent research has supported most of Kounin's findings and has identified a number of additional characteristics of effective classroom managers.

In addition to the characteristics identified by Kounin, effective classroom managers can be described as authoritative rather than authoritarian or laissez-faire in their approach. That is, instead of being overly punitive and controlling or overly permissive, they establish clear rules and high behavioral expectations; enforce rules in a firm, fair, and consistent fashion; and implement a variety of activities that promote student decision making, autonomy, responsibility, motivation, and positive self-concept. In the classroom, they work hard to develop positive student–teacher relations. The critical importance of warm and supportive teacher–student relations is supported by ample research showing that students tend to emulate, listen to, and internalize the values and standards of those adults whom they perceive to be warm, respectful, and caring, as well as research indicating that one of the most common reasons for dropping out of school is the perception that teachers do not care.

Effective classroom managers also understand the critical importance of establishing supportive relations among students, recognizing that students who feel alienated, socially rejected, or lonely are at great risk for a variety of negative social, emotional, and academic outcomes. Thus, effective classroom managers use cooperative learning activities, group decision making, and other techniques to foster feelings of belonging and a "sense of community" in the classroom. In addition to valuing the critical importance of teacher–student and student–student relations, effective classroom managers value the need to develop and maintain supportive teacher–parent relations; thus, they strive to communicate often, and in a positive manner, with the parents of all students.

6. REMEDIATING SERIOUS MISBEHAVIOR

The previously discussed techniques for developing self-discipline and for preventing and correcting misbehavior are sufficient for addressing the needs of the majority of students. To be sure, the number of techniques used, and their intensity and frequency, should vary depending on the needs of individual students as well groups of students. In particular, a combination of the techniques is needed for students considered at risk for future behavior problems. These students are at risk by nature of exhibiting frequent acts of noncompliance at an early age or by nature of certain biological, emotional, or social factors (e.g., attention deficit hyperactivity disorder, a high level of anger, and peer rejection). Research shows that a combination of the previously discussed techniques, especially when implemented in a systematic fashion and over an extended period of time, can be quite effective in reducing negative social and emotional outcomes among at-risk children. An excellent example is the Families and Schools Together program (FAST Track), an empirically supported, federally funded, school-based intervention program for promoting social and emotional competence and preventing future behavior problems. Whereas certain components of the program are provided to all children, other components are provided to those deemed at risk (i.e., those rated by their teachers as being in the top 10% with regard to conduct problems). Beginning in first grade, all children in FAST Track schools receive two or three lessons weekly in social and emotional learning, which are taught by the children's regular classroom teachers. In addition to receiving training in the teaching of lessons, teachers receive training in discipline and classroom management and are provided consultative assistance from support staff, such as a school psychologist or counselor, in managing children with discipline problems. At-risk children participate in a Saturday program that reinforces the skills taught in the social and emotional lessons during the week. Their parents also participate in the Saturday program and receive intensive training in effective parenting. Another important component of FAST Track is the provision of remedial reading instruction to those students reading below grade level, which includes many at-risk children. FAST Track has been found to be effective in preventing aggressive and disruptive behavior and in fostering a positive

school environment in which students are more likely to follow rules, express their feelings appropriately, show greater interest and enthusiasm, and stay more focused and on task. Substantial improvements in a variety of areas were also found among highly disruptive at-risk students. For example, at the end of first grade they were found to have improvement in social problem solving, emotion coping, positive peer interactions, and social acceptance. Perhaps more important, they were rated by parents, teachers, and unbiased observers as less aggressive/disruptive.

Whereas a combination of techniques for promoting self-discipline and for preventing and correcting misbehavior is appropriate for all children, including children viewed as being at risk for future negative outcomes, a small percentage of these students require more intensive or comprehensive services and supports, or what is referred to in this article as remediation of serious or chronic behavior problems. These students are of two general types: (i) those who exhibit serious behavior problems that are harmful to others or to oneself (e.g., violent acts of aggression, drug use, and self-injury) and (ii) those who exhibit less serious behavior problems that may not necessarily concern issues of harm and safety but nevertheless are chronic and interfere substantially with the learning of others and with the development of healthy interpersonal relations with others. These behaviors continue despite the implementation of multiple interventions for developing self-discipline and for preventing and correcting misbehavior. Many, but not all, of these students are the estimated 3–6% of students in need of special education and support services due to an emotional or behavioral disorder. For these students, and others identified as having a disability, federal law [i.e., the Individuals with Disabilities Educational Act (IDEA)] specifies that they can be removed from the regular classroom, without parent permission, if they have committed one of three behavior problems: possession of a weapon, possession of illegal drugs, and any behavior that is judged by a special "hearing officer" as "substantially likely" to injure self or others. IDEA requires that during their removal students with disabilities be provided a continuation of their individualized educational program, including remedial and support services (e.g., counseling).

The codes of conduct in many schools require removal from the regular classroom, or school, as a consequence of serious or chronic behavior problems, as well as for the purpose of protecting the safety of others. Thus, students are often expelled or placed in an "alternative" educational setting—an alternative to the regular classroom and to expulsion. The number of alternative educational settings has increased greatly in recent years, reflecting an increased zero tolerance approach to discipline problems, recognition of the limitations to expulsion, and the fact that IDEA forbids the expulsion of students with disabilities without a continuation of special education services.

Although there is a lack of research on the effectiveness of alterative educational settings, there is considerable research on the effectiveness of various interventions for children with serious and chronic behavior problems. Such research indicates that effective intervention programs have many characteristics in common. They are comprehensive, addressing a combination of multiple risk and protective factors, and broad based, adopting a systems perspective that recognizes the critical importance of schools working closely with families, outside agencies, and communities. Moreover, the interventions are intensive and sustained over time, implemented early, guided by theory and research, and sensitive to developmental differences and the determinants of behavior.

The primary purpose of remedial services and supports is to address the child's current behavior problems and prevent their reoccurrence. Occasionally, such as when serious acts of violence are involved (e.g., use of weapons or a homicide), a crisis exists, consisting of a stress-inducing event in which students, as well as others in the school, experience a breakdown in the ability to cope. In preparation for such unfortunate events, it is recommended that schools have a crisis response team as part of the remedial component to their comprehensive school discipline program. Such a team would consist of specially trained teachers, administrators, and support staff who respond immediately to the crisis and who provide psychological support and resources needed to alleviate existing stress, restore coping skills, and prevent future emotional problems related to the crisis.

Each of the previously discussed four components is a necessary, and critical, part of a comprehensive school discipline plan. Their key features are summarized in Table I. Unfortunately, even among the best schools, correction and remediation are necessary. However, these schools recognize that an emphasis on developing self-discipline and preventing misbehavior with effective classroom management is the most effective way to reduce the need to correct misbehavior and remediate serious behavior problems.

TABLE I
Components of Comprehensive School Discipline[a]

Developing self-discipline
Provide multiple models of desired cognitions (e.g., moral reasoning, social problem solving, and responsible decision making), emotions (e.g., pride, assumption of responsibility over wrongdoing, empathy, and regulation of anger), and prosocial behavior.
Reinforce students, especially with sincere praise, when they exhibit the above qualities.
Implement a research-supported curriculum that directly teaches the above qualities.
Provide ample opportunities for students to apply social and moral problem-solving and decision-making skills to real-life problems and to demonstrate empathy and prosocial behavior.
Correcting misbehavior
Emphasize the use of positive techniques that teach students thoughts, emotions, and behaviors that should replace the given misbehavior.
Never use punishment alone, but always in combination with positive techniques designed to teach appropriate behavior.
If punishment is used, use the mildest technique necessary to improve behavior and a technique that is consistent with the severity of the offense.
Always be aware of the many limitations to the use of punishment.
Preventing misbehavior with effective classroom management
Be authoritative, not authoritarian.
Demonstrate warmth, respect, and caring toward all students.
Establish and maintain close communication with each student's parents or caregivers, and work hard to garner their support.
Provide academic instruction and activities that motivate learning.
Create a physical environment that is conducive to teaching and learning.
Establish clear behavioral expectations, fair rules, and consequences.
Establish predictable procedures and routines.
Frequently monitor student behavior, praising appropriate behavior and responding immediately to signs of misbehavior.
Remediating serious or chronic misbehavior
Provide a comprehensive program that is intensive, sustained over time, and offers a wide range of services, such as adaptive and remedial instruction, consultation to teachers, counseling, and parent education.
Adopt a systems perspective toward understanding, preventing, and treating behavior problems. Whenever possible, always involve parents or caregivers in interventions.
Intervene early: Provide, or arrange for, services when behavior problems first become evident. This may be before or after children enter school.
Be sensitive to developmental and cultural differences in behavior and responsive to the multiple determinants of behavior.
Target multiple risk and protective factors associated with serious and chronic misbehavior, including biological, cognitive, emotional, and environmental factors.
Be aware of special provisions in federal law pertaining to discipline and children with disabilities.
Implement a crisis response plan and team.

[a]Adapted from Bear (2005).

See Also the Following Articles

Emotional and Behavioral Problems, Students with ■ Learning Disabilities ■ School Violence Prevention

Further Reading

Battistich, V., Schaps, E., Watson, M., Solomon, D., & Lewis, C. (2000). Effects of the Child Development Project on students' drug use and other problem behaviors. *Journal of Primary Prevention*, *21*, 75–99.

Bear, G. G. (2005). *Developing self-discipline and preventing and correcting misbehavior*. Boston: Allyn & Bacon.

Bear, G. G., Quinn, M. M., & Burkholder, S. (2001). *Interim alternative education settings for children with disabilities*. Bethesda, MD: National Association of School Psychologists.

Bear, G. G., Webster-Stratton, C., Furlong, M., & Rhee, S. (2000). Preventing aggression and violence.

In K. M. Minke, & G. G. Bear (Eds.), *Preventing school problems—Promoting school success: Strategies and programs that work* (pp. 1–69). Bethesda, MD: National Association of School Psychologists.

Conduct Problems Prevention Research Group. (1999). Initial impact of the Fast Track Prevention Trial of Conduct Problems: II. Classroom effects. *Journal of Consulting and Clinical Psychology*, *67*, 631–647.

Dwyer, K., & Osher, D. (2000). *Safeguarding our children: An action guide*. Washington, DC: U.S. Departments of Education and Justice. [Available at http://www.ed.gov]

Elias, M. J., Zins, J. E., Weissberg, R. P., Frey, K. S., Greenberg, M. T., Haynes, N. M., Kessler, R., Schwab-Stone, M. E., & Shriver, T. P. (1997). *Promoting social and emotional learning: Guidelines for educators*. Alexandria, VA: Association for Supervision and Curriculum Development.

Jones, V., & Jones, L. S. (2004). *Comprehensive classroom management: Creating communities of support and solving problems*. Boston: Allyn & Bacon.

Kauffman, J. M. (2001). *Characteristics of emotional and behavioral disorders of children and youth* (7th ed.). Upper Saddle River, NJ: Merrill Prentice Hall.

McClellan, B. E. (1999). *Moral education in America: Schools and the shaping of character from colonial times to the present*. New York: Teachers College Press.

School Environments

Rotraut Walden
University of Koblenz, Koblenz, Germany

1. Introduction
2. Empirical Evidence Regarding School Environments
3. General Models Describing Environment–Behavior Relations in School Environments
4. Schools of the Future
Further Reading

GLOSSARY

acquisition wall in classrooms Wall that should include the chalkboard and a bulletin board and be in the front of the room; maintenance walls should be along the sides of the room and should include material that helps students to review work; dynamic walls should be at the back of the room and contain student work, school notices, and holiday decorations.

appropriation A process by which the objective environment is transformed into a subjective and personally meaningful environment; constructive appropriations can be expressed through articles such as graffiti, works of art; personalization, changes for use, and the use of an environment. Negative appropriations are shown through vandalism.

centralized schools Schools that are made up of two or three large buildings; decentralized schools are composed of many smaller buildings that are scattered around the campus.

crowding The subjective perception of high density; a greater density of people is linked to negative effects, physiological arousal, (sometimes) an increase in illness, changes in social behavior, a decrease in willingness to help others, and an increase in aggression.

density An objective measurement for spatial demarcation; there is a difference between the amount of available space per person (i.e., spatial density) and the number of people per unit of space (i.e., social density).

open-plan classrooms Large areas of undivided flexible space; the educational changes that accompanied them grew out of the 1960s and were a departure from the standard "egg carton" design used in traditional classrooms.

smart classroom Classroom that includes a computer at the teaching station, the availability of a television or videocassette recorder, and the ability to project or send a computer signal to a large monitor, projector, or each student station; it is a communication-rich classroom that brings the outside world into the traditional classroom.

sociofugal seating arrangements Rows of seats that are turned away from one another, making communication difficult.

soft classroom Classroom with semicircular cushion-covered bench seating, adjustable lighting, a small carpet, and some mobiles; student participation increases markedly in this type of classroom.

stress in schools Conceptualized as an outcome resulting from the lack of congruence between individual needs and goals with the opportunities and constraints afforded by schools.

traditional classroom Classroom that consists of rows and columns of student desks, a horseshoe, or perhaps clusters of four that face the teacher's desk and the blackboard at the front of the room.

The physical school environment can be considered a school's "third teacher," after other children in the class and teachers themselves. The goal of designing new buildings, as well as modifying existing educational facilities, should be the architectural stimulation of achievement, social cooperation, and well-being.

©2004 Elsevier Inc.
All rights reserved.

Achievement and social cooperation are also evaluated in schools by student grades. This article describes the results of studies on different influencing factors: the size of school buildings and classes, the ground layout and façade, the state of the buildings, and their age. It examines the effects of privacy, density, seating positions, and open-plan classrooms. It also reports the results from studies on interior architecture (e.g., access to windows). A number of factors have an immediate effect on the sensory organs: noise, climate, color, lighting, and smell. The article then discusses schoolyards and playgrounds as well as the danger posed by vandalism. It then concerns itself with two conceptual frameworks for analyzing the effects of school environments on the behavior and experiences of teachers and students. A review of recent develops details the criteria that can be used to measure a school's success and details how a school of this kind should be designed. Acquiring a sense of competence with one's environment is deemed to be fundamentally important for future generations, and the opportunities for realizing this are an important part of the physical conditions that a school of the future should offer in the end.

1. INTRODUCTION

After a study carried out by the Organization for Economic Cooperation and Development's (OECD) Program for International Student Assessment (PISA) in 2000, industrialized nations such as the United States and Germany discovered that the performance of their students in relation to international rankings was relatively poor. Since this so-called PISA shock, education is once again being examined closely in various countries. The focus of this has been the development of performance standards, greater control through testing, and a broadening of subjects offered during the afternoons. In relation to this, questions have been raised about the causes of poor performance and discussions have centered on learning conditions and the learning environment. In connection with this, environmental psychologists have also been consulted. The physical school environment can be considered a school's "third teacher," after other children in the class and teachers themselves. Therefore, formal and other architectural concerns of school design should be subordinate to educational demands.

Various studies have shown that the educational demands on schools are multifaceted. Schools should be places that facilitate conflict resolution and encourage learning and living, that is, places to meet for social learning that facilitates conflict resolution and encourages both privacy and public interaction. Learning is defined as a relatively permanent change in behavior as a result of experience. Good teaching can be found in poorly constructed schools, just as bad teaching can exist in well-built schools. What is commonly accepted, however, is that a school building can either facilitate or hinder learning, as observed by Weinstein in 1979.

The effects of the physical setting are moderated by the social and instructional contexts. Schools should involve varied sensory experience and should empower children to interact with their environment in a competent way. They should encourage development of creativity and activities, individual diversity, and teamwork. At the same time, a school can be an effective architectural, aesthetic, ecological, economic, and sustainable role model. Well-designed schools are the prerequisites for the development of successful motivational learning methods that also contribute to enjoyment of learning and a willingness to excel for students and teachers alike. Ultimately, students are the bearers of future innovations for worldwide competition and for society's benefit. The demographic change brought about by an increase in life expectancy demands that fewer and fewer young people support an ever-increasing aging population in industrialized nations.

As early as 1978, Gump calculated that the average person spends approximately 14,000 hours in school from kindergarten through to the 12th grade. Since the results of the international PISA study in many countries were reported, many schools where students were previously taught for only a half-day became full-time schools in the hopes of increasing the achievement levels of students. In Germany, for example, 10,000 of more than 40,000 schools will be full-time by 2007.

As mentioned previously, school environments should help to improve achievement, social cooperation, and well-being. A school of the future should fulfill all of these functions for students, teachers, and parents on a daily basis.

2. EMPIRICAL EVIDENCE REGARDING SCHOOL ENVIRONMENTS

2.1. Special Influences in Schools and Classrooms

During the 1970s and 1980s, most towns in the United States and other countries built very big schools. The

rationale for constructing schools of this size was the expected economic benefits and advantages of offering students a wide curriculum. Today, there is a reversal of that trend, whereby large schools are divided into smaller, more personal units. The goal is, among other things, a greater sense of personalization in the schools that is shown through decorations; participation in school clubs, recesses, and artistic performances; and the expression of a feeling of "I like to be here." The ministries of education had been seeking a solution to poor student performance, problems caused by accidents and crime caused by low levels of personal support, difficulties with supervision, and too much anonymity in schools.

The advantages of smaller schools seem to outweigh those of their larger counterparts. According to Wasley and colleagues in 2000, the benefits of small schools can be found in intimate learning groups, where students are well known by their teachers and can be encouraged to perform better and where adults can look after the students in their care more adequately. Isolation, which manifests itself as alienation, vandalism, theft, and violence, can be avoided. Smaller schools enable children from minority or economically disadvantaged backgrounds to advance sooner, and teachers can be encouraged to draw on their own experiences for the good of the children. In addition, teaching activities can be coordinated more effectively. As early as 1964, Barker and Gump's classic study showed that students in smaller schools are more likely to take part in school activities, even though larger schools offer more opportunities for such activities.

The need for disciplinary measures clearly declined when smaller class groups were introduced in Granby High School in Norfolk, Virginia. In 293 secondary schools in New Jersey, school size was the second strongest predictor (after socioeconomic status) for student achievement in central tests carried out by Fowler and Walberg in 1991. Further positive aspects that accompanied this discovery were greater satisfaction from parents, teachers, students and members of local government as well as a greater responsibility for students on the part of teachers, greater cohesion within groups, and more contact between individuals and the school staff and faculty. Researchers for the Rural School and Community Trust found that smaller schools always outdo their larger counterparts in terms of achievement, particularly with regard to the performance of economically disadvantaged students.

In 1984, Goodlad recommended that elementary schools should not have more than 300 pupils. The highest levels of negative effects were found in high schools with more than 2100 students. Centralized schools with between 500 and 700 students, where children take an active role in activities outside of lessons themselves, and schools with attractive façades appear to be ideal schools for children. In a school of this kind, the number of students is arranged so that there is as little anonymity as possible among students, thereby giving rise to less criminal behavior. Once enrollment exceeds the threshold of 700 students, further increases seem to have little effect. The layout of the school buildings should be centralized to give students as many opportunities as possible to communicate, even between classes. In this way, students can gain more personal recognition through actively taking part in extracurricular activities. In 1982, Glass and colleagues maintained that once the size of a class reaches 20 to 25 students, adding more students seems to make relatively little difference. They concluded that smaller classes lead to better learning environments in nearly every respect, including teacher and student attitudes, interaction, and achievement.

Admittedly, the number of students or the size of a school can only rarely be set because most school buildings are old. Therefore, the teaching staff is asked to make commitments to reduce anonymity among students as much as possible in larger schools by organizing school festivals, letting students share input into the design of classrooms and corridors, and giving students in decentralized schools as much time and as many opportunities as possible to communicate with each other. The question of whether the state of maintenance of a school may have an influence on students was pursued by Gifford in 2002. According to his research, the influence of a school's state on students could not be determined, but he assumed that happier teachers working in more attractive schools are also better teachers. In conclusion, it is clear that students who are taught in new or renovated schools are happier.

The complexity of these concepts renders it difficult to make clear assertions about an optimal room layout. However, some basic assertions should be summarized. Each student needs a certain amount of privacy. Students who show deviant behavior might need a higher level of privacy, and facilities such as "privacy booths" and/or quiet corners could be made available for these children. Whether one decides on measures such as these or similar ways of encouraging privacy, the facilities should be either accessible to all students or not acquired at all. When students sit close to friends, or

near the door, in a classroom with high walls, they feel safe and, therefore, do not need as much personal space among themselves, the teacher, and other students. This can be taken into account by the division of the seating plan in the classroom. In addition, by approaching the students and being closer to them, the teacher can shed some of the distance that his or her status affords, thereby building trust with the students.

An ideal classroom density is between 2.8 and 3.7 square meters of space per student. If the density is higher than this, it can have a negative effect on the students. In any case, investigations carried out on these influences often end up with conflicting results, so that the effects on students must be considered as highly differentiated. No single, all-encompassing assertion about the effects of classroom density on students' performance can be made.

Many investigative results agree that higher classroom density impairs performance when students are carrying out complex tasks, working in groups, or engaging in work that requires mobility. In addition, other influencing factors, such as high levels of motivation, lack of the opportunity to choose one's own seat, a feeling of crowding, and the ages and ethnicities of students, all can exert an influence on students' performance and should be considered. The well-being of students can also be impaired by a classroom density that is too high. Likewise, social behavior gets worse because students must compete for the available teaching materials and the teacher's attention, and this presents itself as an increased level of aggression. This is particularly the case with younger pupils, meaning that a correlation between the cognitive state of students and the reaction to increasing density can be determined.

Just as with classroom density, a student's seating position in a classroom can influence his or her performance, well-being, and social behavior. Therefore, according to results of some research, places in the middle and toward the front of the classroom offer an optimal position for increased achievement and participation in lessons. Furthermore, students connect certain positions within the classroom with emotions and choose their seats with this in mind. In 2001, Bell and colleagues mentioned that students who choose to sit in the middle front seats in a classroom have a higher level of self-confidence than do students who occupy other seats. According to sources in McAndrew's 1993 book, the rear of the classroom is associated with freedom to interact with peers and freedom from teacher control. Consequently, there is a self-selection of pupils in the different rows. However, in Griffith's classic 1921 study, the grades of students who had been assigned to seats alphabetically were carefully recorded over many different classes. Griffith detected that students in the center of the class tended to receive the highest grades, whereas a steady dropoff occurred as the seating position got farther away from the center. If generally quiet students were seated in the "action zone," their participation in classroom discussions did not increase. Communication among students can also be influenced by seating arrangements. Therefore, when regulating the division of seating in a room, as much care as possible should be taken to create a "sociopetal" arrangement that encourages communication, for example, by using group tables.

The open-plan classroom, which found its way into many schools during the 1970s, can have a negative effect on the performance of average students, whereas it encourages those students who are above average. Average students are more susceptible to diversions, whereas those who are above average make use of the diverse learning opportunities. Students' well-being can be affected by an increased level of noise and undesirable distractions. These negative influences are juxtaposed to increased concentration on schoolwork and partially increased communication. However, the open classroom has lessened in popularity over the past 15 years or so due to its negative effects. In 1993, McAndrew concluded that although it is true that most studies comparing the actual achievement of children in open-plan versus traditional classrooms show no major differences, studies that do report differences tend to favor traditional classrooms over open-plan ones. A combination of the traditional classroom and some components of the open-plan classroom could, however, be beneficial for students, and the open concept of group work on tables could be applied. What is important is that the classroom type must correspond to teaching methods, the students' and teachers' need for stimulation or privacy, the type of activity, and the length of time spent on the activity.

It is not only the exterior appearance of a school and its layout that have an influence on students. The interior layout in terms of the division of walls and furnishings, the position of windows, and the type of flooring, decorations, and equipment in the classroom also has an influence. There is much to consider in the interior layout of a school. Temporary walls, or walls that are too low, increase distractibility. The main exterior walls should be high if possible. The division of walls for "acquisition," "maintenance," and "dynamic walls" clearly have advantages if they are used in

classrooms. Empirical results also indicate that decorations and images with "happy" motives should be used.

Having a school that one finds attractive is associated with better achievement, but boys and girls might not find the same decor attractive. In 1990, Cohen and Trostle indicated that girls prefer more windows, color, texture, shapes, and lighting, whereas boys prefer larger settings. When learning occurs in a given classroom, that material is better recalled in the same setting than in a different setting. Some researchers believe that classrooms should have more stimuli and opportunities for environmental exploration that facilitate learning. Others worry that students will have difficulty in concentrating on schoolwork. Both ways depend on the purposes of learning. On the one hand, learning includes social responsibility; on the other hand, it includes acquisition of "dry" intellectual learning material. Bell and colleagues concluded in 2001 that seeking the right fit between person and learning environment is probably the most desirable approach.

Research has shown that the floors should be laid with carpet and that students should always have an unimpaired view of teachers so as to encourage communication between students and teachers. According to research quoted by Gifford in 2002, there is a more direct, close, and learning-oriented contact between students and teachers when the floors are carpeted than when they are not. A lack of dividing walls or screens in classrooms is beneficial so that teachers do not engage in administrative work but instead remain in constant interaction with students. Furthermore, social interaction can be improved using new equipment and facilities. An increase in social communication of this kind is also possible with old furniture if it is renovated.

Some results of research show that the absence of windows leads to increased achievement of students, whereas others demonstrate that it leads to decreased performance, depending on the task given. According to some investigations, a classroom with windows and natural light is a good precondition for high achievement and is especially influential for the positive well-being of students.

Increasing demands are being made on computer equipment and freedom from glare in "smart" classrooms. Furnishings should correspond with ergonomic demands and should be suitable for the size of children and adults.

Despite the fact that "good acoustics" have a strong impact on successful learning in school, the effects of noise are not considered to be important as they should be, according to a 2003 paper by Klatte and colleagues. Deficits regarding room acoustics occur when classrooms are too reverberant, speech intelligibility is too low, and/or background noise levels are too high. Talking and listening to one another is still a major teaching and learning interaction, although there are increasing Internet-based and multimedia forms of learning. Acoustic conditions are often very poor, and verbal communication is only possible by screaming. Physical education teachers often must endure levels of noise between 90 and 100 decibels. The measurement of noise levels in elementary schools in Germany during the entire length of a class period show average levels of 70 to 77 decibels.

In comparison, German legislation concerning workplaces places a limit of 55 decibels on the average noise level of labor in areas primarily devoted to mental activities. If verbal communication is the goal, values of no more than 40 decibels are allowed. Noise is a fundamental stress factor. What is more, the speaker's level should be approximately 10 to 15 decibels higher than the noise level. Teachers who have spent much time in their occupation suffer greatly from noise. The flow of instruction is interrupted by frequent repetitions of information and warnings to be quiet. Reluctance, anger, and fatigue, as well as throat and voice problems, often result. Communication suffers and is reduced. What is more, poor hearing conditions cause verbal information to be misinterpreted or completely misunderstood. This leads to a decrease in resources available for short-term memory and mental processing of aurally given information. A growing number of children suffer from hearing disorders. There are even effects caused by poor acoustic conditions that influence the processing of visually presented information. The students do not adapt to background noise. Despite apparent deterioration in performance with background noise, the noise is often judged to be hardly disturbing or even not disturbing. It can be assumed that learning to read and write is complicated in surroundings that are too noisy. Even though music is associated with being pleasant, the same piece of music played by the same instrument as staccato music instead of legato music distracts from memory performance. (Legato music is characterized by tones attached to each other, whereas staccato chains include short and distinct tones.) The conclusion to findings about noise is not a demand for absolute silence and discipline in classrooms. Free work, singing, and joyful events go along with an increase in noise levels in classrooms. Therefore, it is essential to design classrooms in such a manner that the

potentially negative effects accompanying such positive educational development are prevented. This can sometimes be achieved by simple measures such as renewing felt padding under tables and chairs, checking furniture for squeaky drawers and rattling desk pads, oiling door hinges, and putting up heavy curtains, wide cork boards, and tapestries.

As was shown earlier with noise, a perfect solution to the problem of climate for all students and teachers is hard to find. The guidance levels for the influence of climate on performance of employees in offices are comparable to those of students in classrooms (21 degrees Celsius, 50% humidity, 0.1 meter per second air movement). Air conditioning, often found in the United States, is often set too cold and does not take care of the problem of providing fresh air. Air supply is also dependent on the location of the appliance from which the air is taken in (car park or roof). Gifford concluded that even when care is taken in the control of indoor climates, there must inevitably be a compromise that considers individual differences in amount of clothing, type of activity, and indoor or outdoor movement. There may also be preferred temperature levels based on adaptation to one's home climate or cultural background.

There is no given single color that positively influences all students' achievement, well-being, or social behavior. However, one can determine that well-lit rooms have a more positive effect on students than do sad and dark ones (because well-lit rooms diminish feelings of crowding) and that blue is a better color than red (because blue is more relaxing). Whether or not orange and pink rooms positively influence the social behavior of students cannot be clearly confirmed.

According to Gifford, many individuals prefer incandescent lighting, but it is more expensive than fluorescent lighting. Fluorescent lighting has not been shown to have dramatic negative effects on the performance or health of most students.

From a variety of research on the influence of lighting on performance, well-being, and social behavior of people, it is clear how important daylight is for people, regardless of whether it is streaming into a room directly from a window or is simulated using special lamps. Performance is better, and positive social behavior is more frequent, when daylight is available. However, light that is too bright or strong, as is often found in offices and classrooms, has a negative effect on people's well-being. Cold light increases the negative aspects of women's mood, whereas it decreases those of men, according to a study conducted by Knez in 1995. This gender difference is problematic because it is clear that in a society in which males and females share many spaces, cold or warm light is advantageous for only one of the genders. Therefore, a lighting level must be found that is as good as possible for both genders. A classroom with windows and additional daylight lamps whose brightness can be regulated could be a solution for many students.

Although the connection between fragrances and performance, well-being, and social behavior is described here on the basis of only a single investigation, this is based on results gleaned from previous research that were confirmed or strengthened in this investigation. Therefore, the results may be considered reliable. However, this work did not concern young pupils; rather, it concerned university students. However, the overall positive influence of pleasant fragrances could possibly be transferred to school pupils. A separate investigation would have to be undertaken to substantiate this. Overall, this investigation showed that fragrances can positively influence a person's performance, well-being, and social behavior (the latter in the form of voluntary help). When applied to schools, this means that teachers would use different fragrances according to their requirements. Results from pleasant fragrances are consistent with an interpretation suggesting that pleasant fragrances can enhance task performance by serving as a source of environmentally generated positive affect.

According to McAndrew, schoolyards and playgrounds can be viewed as a type of learning environment if learning about social skills and cognitive abilities takes place there. This is particularly the case with playgrounds because pupils who would otherwise interact in a relatively controlled space must get along with each other independently and with little supervision. What should playgrounds look like to give children the necessary stimulation for learning? They should be made up of different corners that stimulate all of the senses and that are linked to each other using pathways. In addition, playgrounds and schoolyards that are less structured (e.g., adventure playgrounds with mountains of tires piled up) are more stimulating to children's creativity and encourage them to use role-play. Children using playgrounds like these reported that they came up with ever more creative ways of playing with the tires and that they felt a sense of success when they achieved their goals. Children in traditional schoolyards, which are normally fenced in and paved with asphalt, usually play games involving equipment brought from home such as balls of various

sizes. They occupy themselves less with the toys and equipment available because these bore them and can be used in only a limited number of ways. A further article on schoolyards and playgrounds, carried out by Moore in 1989, involved the introduction of nature and natural materials. Through their steady growth and changes, these plants and materials spur the imagination. Children also exercise their muscles by climbing trees. They play with branches and stones and hide in bushes. In addition, these plants and materials bring children closer to nature and let them learn about the problems facing the environment and the ways of working with tools and natural materials. By working with and learning about natural materials, children become more open and competent in interacting with their environment. Further opportunities for improvement of schoolyards and playgrounds include the introduction of safe entrances and exits that can serve as meeting points for children and adults and to bundle social contacts that link people to these places. In this way, greater control over crime can be exercised. In addition, there should be pathways, some of which lead to toilets and telephones and others of which lead to the countryside or surrounding areas.

Goldstein's book on the psychology of vandalism, published in 1996, contained three studies on school vandalism. A study by Pablant and Baxter concerned environmental correlates and confirmed the relationship between vandalism rates and the aesthetic quality and level of preservation of school property. Schools with low vandalism rates were characterized by excellent upkeep of buildings and the surrounding playgrounds. Highly vandalized schools were characterized by neglected maintenance and building upkeep, and the school grounds showed little evidence of aesthetic appeal or efforts to make them pleasant. In general, vandalized schools offered an appearance of abandonment regardless of architectural material or design. The age of the schools was found to bear no relationship to the frequency of vandalism. It was found that the schools' overall quality was generally superior to that of the immediate neighborhood in the low-vandalism group and was generally inferior to that of the immediate neighborhood in the high-vandalism group. Beautification efforts may be more effective in deterring acts of vandalism than are steel fences, electronic sensors, and expensive fortress-like buildings. The latter features did not differentiate between schools with high rates of vandalism and those with low rates of vandalism. Neighborhood diversification of use appears to be an important environmental variable affecting schools' safety. Schools with low rates of vandalism were generally located in areas characterized by multimodality of use such as residences, commercial buildings, and churches. Fully 81% of schools with high rates of vandalism faced uninhabited land such as parks and sports fields. It may be assumed that isolated schools, or schools located in low-density areas, have less neighborhood vigilance and are more vulnerable to vandalism. Schools that recorded lower rates of vandalism could usually be seen by neighbors, were better lit, and had fewer restrictions such as trees blocking the view of the property. Recommendations for decreasing vandalism included encouraging activity on the grounds by opening pathways through school property and by using the community, maximizing the illumination surrounding the school property, and maximizing the aesthetic appeal of the property.

2.2. Environmental Competence as an Educational Goal

Environmental competence is understood to be the ability to cope with the effects of involving oneself in an environment and its appropriation and, thereby, in changing the environment. It involves learning about the environment. According to Gifford, it includes three kinds of learning: (a) personal style, attitudes, and awareness of physical setting; (b) knowledge of physical settings, including technical knowledge, how to unearth new information, knowledge about how social systems control space, and knowledge of person–environment relations; and (c) practical environmental skills such as scouting, matching, personalization, and creative custodianship. Programs in and out of school should teach many different facets of environmental competence, from basic environmental ethics, to environmental protection, to campfire building, to architectural design.

3. GENERAL MODELS DESCRIBING ENVIRONMENT–BEHAVIOR RELATIONS IN SCHOOL ENVIRONMENTS

3.1. A First Taxonomy for the Study of Stress in Schools

Acting on the basis of studies known to that point, Ahrentzen and colleagues in 1982 developed a

taxonomy for the study of stress in elementary school settings. At that time, avoiding stress was of primary concern, and less so was the encouraging performance and creativity using stimulating environmental conditions. Most perceptions were based on the idea that pressure to succeed created learning difficulties. Stress is conceptualized as an outcome resulting from the lack of congruence between individual needs and goals and the opportunities and constraints afforded by the school. Examples of stressful environmental characteristics are "sociofugal" seating arrangements that hinder interaction, whereas "sociopetal" seating arrangements are stressful for students who seek privacy from others during study. Windowless classrooms may produce a decrease in well-being. The amount of aural and visual stimulation in open-plan classrooms may hinder students from performing class activities efficiently and speedily over time. Sometimes, environmental constraints do not influence achievement directly but reduce satisfaction and comfort.

The taxonomy used by Ahrentzen and colleagues arose from the "matching" or "fit" between personal attributes and environmental attributes. This taxonomy includes (a) personal attributes, (b) physical/architectural environmental attributes, (c) social/organizational environmental attributes, and (d) indicators of person–environment fit—all important variables that have been established by the literature.

The personal attributes include past classroom environment experience and the locus of control in open and traditional classrooms. The need for achievement and the related concept of competition/cooperation are meant to interact with environmental characteristics. Children and adults differ in the amount of stimulation desired from the environment. Therefore, the sensation- or arousal-seeking tendency will interact with the environment. Stress will result when there is a mismatch.

Ahrentzen and colleagues differentiated between two kinds of environmental attributes: physical/architectural design and social/organizational elements. Each of these environmental elements can influence educational outcomes directly or indirectly through the other domain. The physical/architectural design features include class enrollment, density, furniture arrangement, seating position, decorations, and aural and visual stimulation. Furthermore, territoriality and boundary clarity, privacy areas, and soft/hard atmosphere of the classroom are mentioned. A soft classroom allows more control by users. Resources mean the accessibility of materials (or teachers). In addition, these authors included openness/containment. They aligned themselves with Moos in their social/organizational elements of the taxonomy to characterize the social climate of a classroom: student involvement, affiliation, teacher support, task orientation, competition, order and organization, rule clarity, teacher control, and innovation. According to Moos's 1979 book, these elements consist of three social climate dimensions: relationship, personal growth, and system maintenance and change. Ahrentzen and colleagues added the teacher–environment fit. The expectations, goals, and needs of a teacher should be supported by his or her environment.

The indicators of person–environment fit are attitudinal indicators such as satisfaction with the classroom setting, commitment to schoolwork, and perceived quality of school life. Behavioral indicators include attendance, disruptive behavior, questioning behaviors, participatory behaviors, persistence in completing tasks, fidgety behavior/nervousness, and creativity. Students will expend more effort on difficult tasks in settings that are supportive of their expectations and needs.

3.2. A Recent Framework for Conceptualizing Person–Environment Relations in Learning Settings

In 2002, Gifford mentioned nearly the same relationships among the personal characteristics of learners, the physical features of the learning setting, and the social/organizational climate of learning-related attitudes and learning behaviors. He drew a model with four boxes and their relationships to each other and called it a framework for conceptualizing person–environment relations in learning settings. When Gifford mentioned (a) the personal characteristics, he also started with school experience and then continued with attitude toward school, motivation to learn, age, gender, and personality. Then, he specified (b) physical features such as size, noise level, climate, population density, and design. The literature offers evidence that supports adding fragrances/odors to the list. By (c) the social/organizational climate rules, he meant curriculum and teaching styles, regardless of whether they are progressive or traditional in orientation. These all influence (d) reactions such as learning-related attitudes (e.g., satisfaction with the school or classroom, the desire to learn) and reactions such as learning behaviors (e.g., on-task time, class participation, questioning, appropriate

or inappropriate activity, persistence, creativity, learning, performance).

3.3. Newer Methods for Assessing School Buildings and Criteria for Improvements in the School Environment

In 2002, Henry Sanoff advocated the participation of the local community, students, parents, and teachers in the design and construction process, and in 2001 he developed methods for assessing school buildings. One of these is the six-factor school assessment scale. Furthermore, Sanoff proposed an observation form and rating scales for school buildings, encompassing everything from indoor and outdoor learning space, to informal social areas, to the dining space, to classrooms. He also used open-ended questions such as "I wish my school . . ." and used photographs of the school to assess it.

The school environment should contribute to the overall success of schools. According to Duke and Trautvetter in 2001, partially measurable objectives for school design include, among other things, how well they contribute to a more caring and supportive school culture and to improved achievement in standardized tests and better grades overall. The appropriate social, administrative, educational, or disciplinary measures can bring about reduced disciplinary referrals, suspensions, and exclusions from lessons. These measures can also facilitate faster intervention for children with learning difficulties and greater participation in different kinds of lessons and other instructional activities. Finally, parental involvement and better communication between the school and home can be increased. In the long term, successful schools will see fewer teachers taking sick leave or being lost through early retirement.

4. SCHOOLS OF THE FUTURE

How should schools of the future be built from the point of view of environmental psychology? In a qualitative study conducted by Walden and Borrelbach in 2002, the architects of six innovative schools, chosen by the editorial journalists of professional trade journals, answered this question in interviews. Intelligently designed schools should offer opportunities for users to make their own adjustments to the environment in response to their own levels of arousal, stress, overload, fatigue, and attitudes. Such adjustments include physical conditions such as lighting, color, acoustics, noise, heating, air-conditioning, and the ergonomic composition of the buildings and their furnishings. Users' identification with school buildings is facilitated through participation and appropriation processes. This sense of control by the users should result in increased responsibility for damage and neglect as well as in a decline in vandalism. In Germany, Walden used data from supervisory school authorities in medium-sized cities to estimate the cost of vandalism to be approximately 50 million Euros per year for more than 40,000 schools. (In 1990, Stoner and colleagues estimated that the annual monetary cost of school vandalism in approximately 84,000 schools in the United States amounted to approximately $600 million, according to Goldstein's book from 1996.)

The special needs of children and teenagers should be considered, for example, in the design of playgrounds, entrances, and circulation areas. Specific educational areas should be provided for learning, working, discovery, play, orientation, physical education, and recreation. Groups with special needs and disabilities would benefit not only from legally required entrances with wheelchair access, widened doors, walking frames, steps, handrails (for use by children as well), elevators, and rest rooms but also from specially designed play items.

What are the worst mistakes that one can make when building a school?

- Building too big (results in anonymity)
- Building too small (e.g., few provisions for clubs and activities, lack of specialized rooms)
- Poor connections between entrance and classrooms
- Waste of energy (due to drafts and bad heating insulation)
- Classrooms too small
- Poor arrangement of facilities for faculty, staff, and students as well as for the community
- Bad noise insulation
- Insufficient facilities for the disabled
- Poor upkeep and maintenance (leads to an increase in vandalism and uncleanliness)

The school of the future should offer a quality of experience, even at the planning stage, in consideration of the wishes of students and teachers. Also, it should provide a livable design and exude a pleasant atmosphere. To achieve a feeling of "being at home" in a school of this kind, or one where students and teachers spend a lot of time, it is crucial that children, teachers, and parents are able to participate in its

design. A further effect of "nice" or "pleasant" self-design of schools is an increased feeling of responsibility for one's environment, an increase in creativity and environment intelligence, and a decrease in vandalism (Figs. 1–6).



FIGURE 1 Provisions for rich sensory experiences and privacy (school by Peter Busmann). Reprinted courtesy of Asanger Press, Germany (Walden & Borrelbach, 2002).



FIGURE 2 Sound Stone, a stone that buzzes. Reprinted courtesy of Asanger Press, Germany (Walden & Borrelbach, 2002).



FIGURE 3 Classroom with decorations. Reprinted courtesy of Asanger Press, Germany (Walden & Borrelbach, 2002).



FIGURE 4 Wood as warm material in schools (by Peter Hübner). Reprinted courtesy of Asanger Press, Germany (Walden & Borrelbach, 2002).

Schools looking to the future should take the following into account:

- The right of all users to participate in decision making
- The use of environmentally friendly construction materials
- The opportunity to regulate environmental conditions (e.g., lighting, heating) oneself while at the same time using sensory technology
- The importance of a clear orientation, beginning with the entrance
- The importance of natural lighting
- The flexibility and multifunctionality of rooms
- The arrangement of classrooms and hallways that offer opportunities for privacy
- The importance of providing opportunities for teachers' privacy
- The value and importance of acoustics and noise prevention

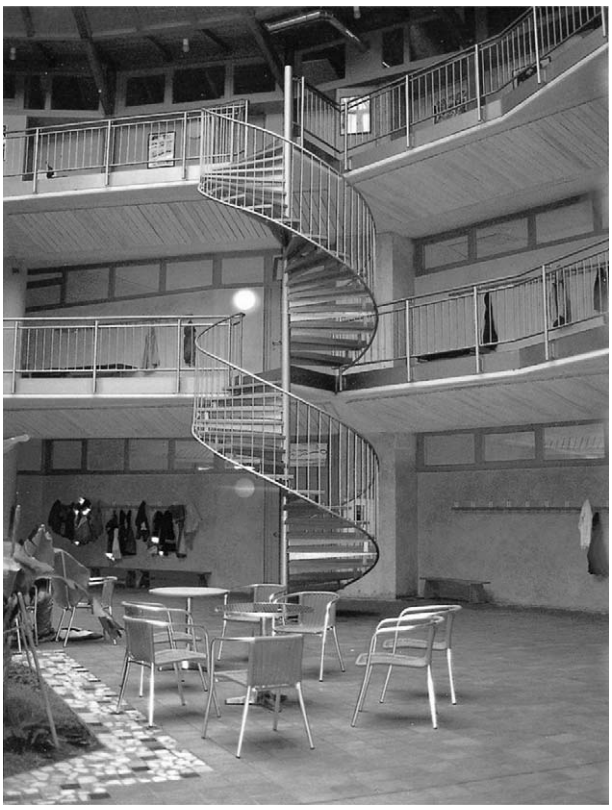


FIGURE 5 Oasis (with plants) for meeting opportunities (by Peter Hübner). Reprinted courtesy of Asanger Press, Germany (Walden & Borrelbach, 2002).



FIGURE 6 Welcoming entrance to a school (by Friedensreich Hundertwasser). Reprinted courtesy of Asanger Press, Germany (Walden & Borrelbach, 2002).

- The importance of providing safe, accident-proof toys and equipment in the schoolyard
- The importance of providing rich "sensory experiences"

But who will pay for such ideal schools? This objection can be met, for example, with reference to advanced methods of heat insulation, energy conservation, and cost-effective building construction. Above all, however, a building that is perceived by its occupants as being "theirs" is more likely to be protected against vandalism and disrepair and, therefore, may become more affordable in the long run. As a consequence, a "well-being school" does not necessarily have to be expensive.

See Also the Following Articles

Environmental Assessment ■ Environmental Design and Planning, Public Participation in ■ Environmental Stress ■ Full-Service Schools ■ Person–Environment Fit

Further Reading

Ahrentzen, S., Jue, G. M., Skorpanich, M. A., & Evans, G. W. (1982). School environments and stress. In G. W. Evans (Ed.), *Environmental stress* (pp. 224–255). New York: Cambridge University Press.

Barker, R. G., & Gump, P. V. (1964). *Big school, small school*. Stanford, CA: Stanford University Press.

Bell, P. A., Greene, T. C., Fisher, J. D., & Baum, A. (2001). *Environmental psychology* (5th ed.). Fort Worth, TX: Harcourt Brace.

Duke, D. L., & Trautvetter, S. (2001). *Reducing the negative effects of large schools*. Washington, DC: National Clearinghouse for Educational Facilities.

Gifford, R. (2002). Educational environmental psychology. In *Environmental psychology: Principles and practice* (pp. 296–336). Colville, WA: Optimal Environments.

Goldstein, A. P. (Ed.). (1996). *The psychology of vandalism*. New York: Plenum.

Gump, P. V. (1991). School and classroom environments. In D. Stokols, & I. Altman (Eds.), *Handbook of environmental psychology* (pp. 691–732). Malabar, FL: Krieger.

Klatte, M., Meis, M., Nocke, C., & Schick, A. (2003). Classroom acoustics: Acoustical conditions in schools and their impact on pupils. In M. Cudina (Ed.), *Proceedings of the First Congress of Alps Adria Acoustics Association and Third Congress of Slovenian Acoustical Society* (pp. 9–16). Ljubljana, Slovenia: Slovenian Acoustical Society.

McAndrew, F. T. (1993). *Environmental psychology*. Pacific Grove, CA: Brooks/Cole.

Moore, R. C. (1989). Playgrounds at the crossroads: Policy and action research needed to ensure a viable future for public playgrounds in the United States. In I. Altman, & E. H. Zube (Eds.), *Public places and spaces* (Vol. 10, pp. 83–120). New York: Plenum.

Moos, R. F. (1979). *Evaluating education environments*. San Francisco: Jossey–Bass.

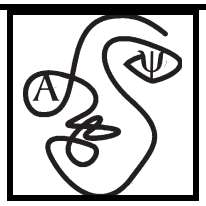
Pablant, P., & Baxter, J. C. (1996). Environmental correlates of school vandalism. In A. P. Goldstein (Ed.), *The psychology of vandalism* (pp. 213–233). New York: Plenum. (Original work published 1975)

Sanoff, H. (2001). *School building assessment methods*. Washington, DC: National Clearinghouse for Educational Facilities.

Sommer, R., & Olsen, H. (1980). The soft classroom. *Environment and Behavior*, *12*, 3–16.

Walden, R., & Borrelbach, S. (2002). *Schulen der Zukunft: Gestaltungsvorschläge der Architekturpsychologie* [Schools of the future: Design proposals from architectural psychology]. Heidelberg, Germany: Asanger.

Weinstein, C. S. (1979). The physical environment of the school: A review of the research. *Review of Educational Research*, *49*, 577–610.



School Psychology, Overview

Howard M. Knoff

Arkansas Department of Education, Little Rock, Arkansas, USA

1. Qualifying School Psychology as a Profession
2. The Elements of an Effective School and the Contributions of Applied Psychology to Effectiveness
3. What School Psychologists Do to Be Effective
4. This Section of the Encyclopedia
5. Conclusion

Further Reading

GLOSSARY

credentialing and licensure standards Standards that typically are the state-specific training, supervised practice, and continuing education criteria that result in a school psychologist being certified for independent practice either in the schools (credentialing) or in a community-based or private practice setting (licensure).

functional assessment Assessment method that focuses on either increasing or establishing appropriate behavior or decreasing or eliminating inappropriate behavior through a process that evaluates students' strengths and assets along with their weaknesses and deficits. The process involves using "authentic" or curriculum-based assessments that directly evaluate the skill sets or domains of interest that determine why students are or are not making progress in various areas, and then linking the assessment results directly to strategic interventions to address targeted problems or concerns.

"No Child Left Behind" Act (NCLB) and Individuals with Disabilities Education Act (IDEA) legislations These two federal laws guide much of general education and special education practices in schools and school districts. The NCLB is the current elementary and secondary education law for the United States that guides expectations for students' academic progress, teacher quality, positive school climate and safety, and other annually evaluated outcomes. The IDEA guides and monitors the delivery of services that are provided to special education students who are identified as having one of 13 specific disabilities, requiring that they receive a free and appropriate public education (FAPE), an individualized education program, and other due process and educational rights.

professional practice and ethical standards in school psychology Standards that govern the professional practices and interactions of school psychologists as they (a) provide assessment, consultation, intervention, and evaluation services to students and parents, staff and administrators, and schools and other organizations; (b) interact with colleagues and other professionals; and (c) work in service delivery units in the field and/or school psychology training programs at the university level. The ultimate goal of these standards is the protection of the rights and integrity of all individuals and parties who are directly or indirectly impacted by provided services.

school psychology One of the four original specialty areas of professional psychology recognized by the American Psychological Association—specifically, organizational psychology, clinical psychology, counseling psychology, and school psychology. While there is a common science of psychology undergirding all four of these professional areas, school psychology focuses on the application of social and organizational, educational and learning, child and adolescent, normal and abnormal, and biological and ecological psychology to the school and schooling process.

scope of school psychological services School psychology services are provided to all students along a continuum, from (a) those who benefit from effective instruction without

the need for additional support services, to (b) those who need strategic interventions in order to make appropriate academic, social, emotional, and behavioral progress, to (c) those who need intensive or crisis management services in order to be successful.

School psychology is one of the four original specialty areas of professional psychology recognized by the American Psychological Association—specifically, organizational psychology, clinical psychology, counseling psychology, and school psychology. While there is a common science of psychology undergirding all four of these professional areas, school psychology focuses on the application of social and organizational, educational and learning, child and adolescent, normal and abnormal, and biological and ecological psychology to the school and schooling process. More specifically, school psychology works from a systems level to an individual level. At the systems level, it emphasizes the importance of creating safe and positive school environments that facilitate effective instructional processes that result in students' academic, social, emotional, and behavioral progress and success. At the individual level, it emphasizes the importance of guiding, intervening with, and evaluating students' health and mental health, and ways to help them develop and respond to protective factors such that they become independent learners, effective behavioral self-managers, and resilient in the face of life's passages and crises.

1. QUALIFYING SCHOOL PSYCHOLOGY AS A PROFESSION

School psychology is a profession. Guided by two national professional associations, the National Association of School Psychologists and the American Psychological Association's School Psychology Division, school psychology has training standards, professional practice and ethical standards, credentialing and licensure standards, and continuing education expectations.

School psychology is a post-baccalaureate profession. In most states, school psychologists are certified by their state departments of education or psychological licensure boards for practice in the schools at the educational specialist (Ed.S.) level. Typically, this involves 60 graduate hours of defined coursework and a school-based or school-related internship. Some school psychologists earn the doctorate degree (Ph.D., Ed.D., or Psy.D.). Typically, this involves a minimum of three years of practice, a full year internship, a specialization area, and research culminating in a dissertation. Currently, about 75% of the profession practices at the Ed.S. level, and 25% practices at the doctoral level.

Relative to course work, school psychologists receive university training, clinical practice, and supervision in the following domains: (1) data-based decision making and accountability; (2) interpersonal communication, collaboration, and consultation; (3) effective instruction and development of cognitive/academic skills; (4) socialization and development of life competencies; (5) student diversity in development and learning; (6) school structure, organization, and climate; (7) prevention, wellness promotion, and crisis intervention; (8) home/school/community collaboration; (9) research and program evaluation; and (10) legal, ethical practice, and professional development. Demonstrating their adherence to a published set of standards and criteria, university training programs are either approved (at the doctoral and Ed.S. levels by the National Association of School Psychologists) or accredited (at the doctoral level by the American Psychological Association), thus providing an independent and objective level of accountability and quality control.

The practice of school psychology in the field is guided by professional practice and ethical standards. Professional practice standards, largely generated by the two national professional associations and state certification and licensure entities, govern the professional practices and interactions of school psychologists as they provide assessment, consultation, intervention, and evaluation services to students and parents, staff and administrators, and schools and other organizations. These activities include professional practices in the 10 training and practice areas listed above, but they also involve the professional interactions of school psychologists with the clients, the consultees, and others that they serve, and how they carry out their various responsibilities. In a broader sense, professional practice standards also address how school psychology units or departments (often in larger school districts in which multiple psychologists are employed to provide a full continuum of comprehensive services) should function to ensure appropriate and responsible practice.

Critically, it should be noted that school psychologists practice in a number of different settings. As they largely focus on the academic and social, emotional, and behavioral development of children

(starting generally at the preschool level) and adolescents (sometimes, for students with disabilities, through age 21), school psychologists are employed in schools; community mental health and other agencies; alternative education, juvenile justice, and prison settings; hospital and psychiatric treatment facilities; private practice; universities; and other specialized settings. In many of these settings, they interact not only with psychologists from other specialty areas, but also with teachers and administrators, social workers and counselors, speech and occupational/physical therapists, physicians and psychiatrists, and other health and mental health professionals.

Ethical practice standards, also generated by national associations and state entities, govern the professional practices and interactions of school psychologists relative to protecting the rights and integrity of all individuals and parties who are directly or indirectly impacted by provided services. Involving morally and legally responsible behavior, ethical practice standards are the "internal" guidelines that define, for individual psychologists and individual practice, appropriate professional scope, boundaries, communications, representations, interactions, and services. While addressing personal behavior, ethical practice standards are reinforced, at the state and national levels, by colleagues appointed to boards of ethical practice. When allegations of ethical violations are forwarded to these boards, they objectively gather the data and information necessary to make a judgment on the behavior, either dismissing the allegations as unfounded or requiring restitution and/or sanctions when the allegations are validated.

In professional practice, the most important ethical practice standard involves the need for school (and other) psychologists to practice within the scope of their training and expertise. Acknowledging that no one professional can be expert in all possible areas of practice, this standard ensures that psychologists provide services only in areas in which they can provide appropriate and effective services—independently or under supervision. For school (and other) psychologists, this is the ultimate protection for the consumer, a protection that is essential given the sometimes tenuous nature and enormous responsibility for safeguarding the psychological health and development of children, adolescents, and their families.

Credentialing and licensure standards are typically state specific, and they are written and enforced by a state's Department of Education Certification Board and Department of Professional Practice, respectively. Most often, these standards are codified in law and/or regulation, as both of these departments must submit them to their state legislatures for approval, and they require the psychologist to pass some type of written national or state exam.

The school psychology credential is the Department of Education's official certification that an individual has received the appropriate university course work, training, experience, and supervision to practice in the schools as an independent practitioner. Often, school psychologists receive an initial or provisional credential, which is then upgraded to a full or independent practice credential after two or three years of full-time supervised practice in the field.

Licensure, meanwhile, is the Board of Psychology's (or the equivalent) certification that an individual has received the appropriate university course work, training, experience, and supervision to practice in the community as a private, independent practitioner. Although licensure is most often conferred only at the doctoral level, there are currently more than 10 states in which a school psychologist can be licensed at the non-doctoral level. Whereas eligibility for school psychology credentialing generally requires university-based supervision, eligibility for licensure typically requires both university-based supervision and post-graduation supervision of practice in the field by a licensed psychologist. Licensure ultimately allows school psychologists to practice in virtually all settings—public and private—while credentialing generally limits the scope of practice to the schools. Some states, however, issue only a school psychology license, and some states require school psychologists to have a credential (even if they also have a license) to do school-based work.

If this were not confusing enough, the National Association of School Psychologists has issued national certification in school psychology (NCSP) to qualified school psychologists since 1988. Through their state credentialing regulations and procedures, approximately 30 states now recognize the NCSP for immediate credentialing. The NCSP has not only established a more uniform set of training standards for entry into the field, but it has also facilitated a level of reciprocity so that school psychologists can more easily move into new states and begin to practice professionally.

School psychology's professional practice and ethical standards, and most state credentials or licenses, require school psychologists to maintain their clinical skills through continuing education. Typically, this involves taking a certain number of courses, spending a specific number of hours at workshops or in-service, and/or

engaging in documented types of professional activities or experiences over a defined period of time. The goal is to ensure that practitioners continue to have the most up-to-date training, experience, and professional standing in a field that is constantly changing and upgrading. This is essential so that they have the knowledge, skill, and confidence to make effective and evidence-based professional decisions over time, and to implement effective interventions and efficient services.

2. THE ELEMENTS OF AN EFFECTIVE SCHOOL AND THE CONTRIBUTIONS OF APPLIED PSYCHOLOGY TO EFFECTIVENESS

Over the past two decades, there have been increasing expectations that schools and school districts demonstrate that all students are successfully meeting specific academic standards and outcomes on an annual basis. These expectations have been federally legislated through such laws as the "No Child Left Behind" Act (NCLB), which is the current elementary and secondary education law for the country. The NCLB is reinforced by the federal IDEA (Individuals with Disabilities Education Act) legislation. IDEA guides and monitors the delivery of services that are provided to students who are identified with one of 13 specific disabilities (e.g., specific learning disabilities, cognitive or intellectual disabilities, emotional or behavioral disturbances, deaf/hard of hearing or blind/visual impairment disabilities) in a number of specific outcome areas: the number of students with disabilities graduating from high school with a regular diploma versus dropping out with no diploma at all; the disproportionate representation of minorities identified in certain disability categories or recommended for school suspension or expulsion due to alleged disciplinary actions; the participation of students with disabilities in their districts' high-stakes or state standards assessments and their demonstrated adequate yearly academic progress on those tests; and the percent of students with disabilities who are taught in general education classrooms for the majority of the school day. Beyond these focus areas, schools are now also mandated to provide a free and appropriate public education to all students; to demonstrate that they have positive, safe, and supportive school and classroom climates; to employ "qualified teachers" in all classrooms; to notify parents annually as to the data-based effectiveness of their children's schools; and to track and maintain high standards for all students regardless of gender, socioeconomic background, race, culture, language of origin, or perceived potential.

To accomplish all of these requirements, schools must identify, organize, and implement evidence- and research-based procedures that facilitate school effectiveness, and then engage in a process of formative and summative evaluation that results in continuous improvement. Existing research can be merged into an "effective schools blueprint" that specifies the characteristics and variables that predict students' academic success. In Arkansas, for example, this blueprint has been organized with the following 11 primary standards:

Standard 1: Effective schools have committees and/or teams that support their professional development, curriculum and instruction, and parent and community outreach activities and school goals. Effective schools also have committees and/or teams that support students' successful attainment of specific academic and social, emotional, and behavioral goals.

Standard 2: Effective schools have an accessible and public document that outlines the scope and sequence of all goals and objectives in all curricular areas taught in the school. This scope and sequence document is cross-referenced with state standards and benchmarks, and it is used as a formative evaluation tool to track student progress. In academic areas, the classroom-based use of curriculum-based assessment and/or measurement is used to track students' progress along these scope and sequence goals and objectives.

Standard 3: Effective schools have a plan and implement a school-wide positive behavioral support system that includes: (a) social skills instruction for all students, and (b) a developmentally successful, school-wide accountability system that includes behavioral expectations connected to incentives and positive responses, and levels of inappropriate behavior connected to negative responses, consequences, or strategic interventions. Included in this plan are prevention and intervention approaches that ensure positive student behavior and safety in common areas of the school (e.g., hallways, bathrooms, buses, cafeteria, playgrounds) and that address potential and actual incidents of teasing, taunting, bullying, harassment, and physical aggression. Also included in this plan are periodic evaluations of school climate and student, staff, and parent satisfaction with the school-wide system.

Standard 4: Staff in effective schools receive ongoing training, evaluation, feedback, and supervision (when needed) in the development and implementation of (a) effective instructional lessons, (b) approaches that effectively link students' acquisition of skills to curricula and instruction, (c) early academic interventions for students when needed, and (d) more intensive classroom-based interventions that are recommended through data-based problem solving and consultation processes. All of these processes result in demonstrated student progress and success.

Standard 5: Staff in effective schools receive ongoing training, guided discussions, supervision (when needed), and follow-up in effective classroom organization and behavior management approaches, and in other approaches that maximize students' time on task, their academic engagement, and their effective use of allocated academic learning time. Staff also use effective and flexible grouping patterns for their students to maximize learning outcomes, and they create and maintain positive, safe, and cooperative learning environments.

Standard 6: Staff in effective schools receive ongoing training, evaluation, feedback, and supervision (when needed) in data-based problem solving and the functional assessment of students' academic and behavioral skill and progress. This training includes the analysis of student records and cumulative folder information, and a linking process between problem analysis and intervention. This training also is especially focused on students who need strategic and/or intensive interventions due to their difficulties in acquiring or demonstrating such skills taught through more routine instruction.

Standard 7: After completing functional assessments, staff in effective schools have the skills (or have access to consultation that can provide the skills) to implement more intensive behavioral and academic/instructional interventions. These intensive interventions are specified for all staff, and they receive ongoing training, evaluation, feedback, and supervision in their implementation.

Standard 8: Effective schools have a written, systematic problem-solving and functional assessment process that involves grade, department, and building-level teams that provide pre-referral interventions for students who are not making sufficient academic and/or social, emotional, or behavioral progress. Related service and other professionals (e.g., school psychologists, speech pathologists, social workers) are represented on these teams, and they provide consultation and intervention support to general education teachers as a first step in the pre-referral process.

Standard 9: Effective schools have an organized, formal, and ongoing process to articulate (or transition) students, academically and behaviorally, from grade to grade and teacher to teacher at the end of the year and during other school year transitions. This articulation process includes strategic and intensive needs interventions that have been developed and implemented for students who are not making sufficient academic and/or social, emotional, or behavioral progress. It also includes students' receiving special education, vocational, and other services. Finally, it includes students' transitioning into the workplace or into alternative (or similar) settings.

Standard 10: Effective schools have a written and systematically implemented parent and community outreach program. The parent outreach program includes activities that encourage parent participation in school activities and help parents to understand the school's goals, objectives, programs, and desired student outcomes. The community outreach program includes collaboration with social service, mental health, law enforcement, and other relevant agencies such that there is direct and indirect support for all school and schooling goals and objectives.

Standard 11: Effective schools have a data-management system and evaluation process, which is supported by computer technology whenever possible, to formatively and summatively track the school's progress toward its stated school, staff, parent, and student goals. For students receiving strategic or intensive services, this data base should be able to track the implementation of student academic and/or behavioral intervention plans as written.

Clearly, the 10 domains that form the foundation of school psychology are well represented within these standards, as are many of the applied areas of psychology (i.e., social and organizational, educational and learning, child and adolescent, normal and abnormal, and biological and ecological psychology) that contribute to the effectiveness of a school. Cross-referencing the 10 domains and the 11 standards defining effective schools, however, suggests a prototype representing a "scope and sequence" of school psychological service delivery. Specifically, the scope of school psychological services involves all school-aged students—typical students, students who need strategic interventions, and students who need intensive services. The sequence of these services involves data-based problem solving that includes functional assessment, consultation, and interventions.

While there often appears to be a dichotomy between general and special education services, school psychology views students along a continuum, from (1) those who benefit from effective instruction without the need for additional support services to (2) those who need strategic interventions in order to make appropriate academic, social, emotional, and behavioral progress, to (3) those who need intensive or crisis management services in order to be successful. For the first group, there is an assumption that effective instruction is occurring in the classroom, and that students are acquiring new skills on the way toward becoming independent learners (academically) and effective self-managers (behaviorally). School psychologists focus predominantly on primary prevention—helping schools to develop and maintain positive behavioral support systems, helping teachers to match effective instructional approaches delivered using evidence-based curricula to individual students and student group learning styles, and helping students to adopt and use healthy behavioral styles and patterns.

The strategic interventions group of students does not acquire skills as quickly or as easily as those in the first group, but with the right assistance and interventions, they can be successful. In order to know how to strategically intervene on behalf of these students, it is critical that school psychologists and other educators determine why they are having difficulty. This involves data-based functional assessment, because there can be a number of different explanations for a student's lack of progress and, thus, a number of different interventions. In brief, school psychologists, working here at the secondary prevention level, often organize their functional assessments by considering six important domains: the student, the teacher or instruction, the curriculum, the classroom (which includes the peer group), the school and district, and the home and community. The first three of these domains generally explain students' difficulties in causally related ways, while the latter three domains provide explanations that are correlationally related. Causally related factors (e.g., poor student motivation, ineffective instruction, inappropriate student–curriculum match) directly explain what has caused an existing skill deficit and exactly what interventions need to be implemented. Correlationally related factors (e.g., a peer group that ridicules academic achievement, large class sizes, poverty at home) relate to the target problems, but do not directly cause them. For example, while poverty may contribute to a student's lack of reading experience and working vocabulary, it is these two latter factors that actually cause the problem.

Functional assessments, to the greatest degree possible, occur in the settings in which the desired or problematic behaviors are occurring. While focused on either increasing or establishing appropriate behavior or decreasing or eliminating inappropriate behavior, functional assessments look at both student strengths and assets along with student weaknesses and deficits. In addition, they also involve "authentic" or curriculum-based assessments that directly evaluate the skill sets or domains of interest (e.g., improved performance in the area of mathematical computations are assessed using math problems that are sampled from the textbook being used in the classroom or the standards being assessed on a state standards test). Finally, functional assessments must provide results that are "psychometrically sound"—that is, that are both reliable and valid so that appropriate interventions can follow. Ultimately, functional assessment results need to link directly to strategic interventions. In this way, interventions with a high probability of success that address the root or source of the problem will hopefully be implemented with successful outcomes.

Finally, intensive needs or crisis management students are not successful in typical school settings and environments without significant levels of support and intervention. Although the same six functional assessment domains should be investigated to determine their contributions to the existing challenges, specific students' characteristics often are the most directly relevant. Indeed, many of these most-challenging students have disabilities that impair their academic and/or social, emotional, or behavioral progress. Others have mental illnesses that similarly contribute. For these students, school psychologists typically work at the tertiary prevention level as part of a school-based multidisciplinary team or community-based interdisciplinary team, completing ongoing functional assessments that lead to comprehensive, multidimensional "wrap-around" or multisetting, 24-hour interventions. Ultimately, these interventions serve to support the student's progress in the least restrictive environment possible and in the least intrusive way, while also helping the student to avoid future crisis situations or, at least, to minimize their impact and frequency over time. Intensive needs students require a great deal of time, effort, coordination, and resources when they are present in schools. Often, schools must accommodate for or help these students to compensate for their

challenges—sometimes in different settings or alternative programs. At other times, schools must recognize that these students will make progress only at their own pace and in their own way. They can learn, but in ways that are more idiosyncratic than their typical peers.

As noted, school psychological services involve data-based problem solving that includes functional assessment, consultation, and interventions. This approach contrasts with the more traditional "refer, test, diagnose, and place" mode of operation. In this more traditional approach, school psychologists worked primarily at the secondary and tertiary prevention levels, intervening with students with pre-existing or emerging difficulties. While doing this, their assessments focused predominantly on the students, working to diagnose and label their deficits or pathologies so that they could be placed in appropriate special or therapeutic settings. Some of the limitations of this approach included (1) the absence of multisetting, multisource, and multi-instrument ecological assessments; (2) the assumption that a diagnosis or label leads to functional, appropriate intervention strategies and plans; and (3) the belief that where an intervention occurred was more important than what actually occurred.

Functional assessment, as described earlier, exists within the data-based problem-solving context involving problem identification, problem analysis, intervention, and evaluation. Completed to determine why a problem situation exists, functional assessment links directly to interventions that have the highest probability of increasing or decreasing specific behaviors or interactional patterns. In the more contemporary approach to school psychological services, what needs to be implemented (i.e., the interventions) is determined first. Where the services need to delivered (i.e., the setting) is identified next, representing the best setting both for effective service delivery and to obtain the fastest results with the greatest degree of change. This more contemporary approach also focuses its attention on student outcomes and not student diagnosis or placement; its interventions focus on behaviors and not labels.

The completion of a functional assessment, with its associated intervention recommendations, however, does not ensure effective intervention. This is where consultation comes in. Consultation is a collaborative process whereby school psychologists work directly with "consultees"—teachers, parents, counselors, and others who have the primary responsibility for implementing recommended interventions. Through the consultation process, school psychologists ensure that interventions are acceptable to consultees before they are attempted, that they have a high probability of effecting meaningful change or outcomes, that they are implemented with integrity, and that they can be successfully generalized to different people, settings, times, and situations. In working with school-based consultees, school psychologists have two goals: (1) to help to "solve" or remediate the referred problem situations, and (2) to enhance the skills of the consultees such that they are able to implement the interventions developed more independently over time—both with the original students of concern, and, ideally, for other students with similar concerns in the future.

Critically, consultation is not limited to students with academic or behavioral challenges. Paralleling the foundational areas of psychology, there are at least seven different areas of psychological consultation: organizational consultation, (group) process consultation, instructional consultation, behavioral consultation, mental health consultation, advocacy consultation, and community-based strategic consultation. Recognizing that school psychologists deliver services at the home and community, district and systems, team and grade, and peer group and individual student levels, it is possible that any of these consultation areas might be practiced.

Finally, as previously noted, successful school psychological services should be defined by successfully implemented interventions that result in appropriate student progress and outcomes. At the same time, as with the different areas of consultation, some interventions may focus on the home and community, the school district and its schools, the staff and the curriculum, as well as the peer group and its individual students. Thus, interventions that result in effective schools are as important as, and related to, effective interventions for and with individual students.

3. WHAT SCHOOL PSYCHOLOGISTS DO TO BE EFFECTIVE

While an understanding of successful school psychological services is important, these services require the training, supervision, and practice of effective school psychologists. Within the context of consultation, argued here to be school psychologists' primary role in the schools, the characteristics of effective school psychological consultants have been empirically derived. In short, effective consultants possess four distinct sets (or factors) of related characteristics and/or skills: they have effective

interpersonal skills, sound problem-solving skills, and good consultation process skills, and they practice in ethically and professionally appropriate ways. While the importance of the latter three areas have been discussed, it is significant that the interpersonal skill area represent the strongest of these four factors, accounting for the greatest amount of the variance. This is not surprising from a psychological perspective—whether organizational, social psychological, or even therapeutic. Indeed, without people who develop relationships in which they communicate productively, work collaboratively, demonstrate care and empathy consistently, and show a commitment to desired outcomes, problem solving will not be successful, consultation will not be acceptable, interventions will not be implemented with integrity, and schools will not be effective. This means that school psychologists recognize and attend to the fact that effectiveness involves both process and content. The process involves, interpersonally, positive interactions with others; the content involves, professionally, appropriate information, strategies, and interventions that are collectively discussed and carefully implemented.

But school psychologists need more than consultation skills to be effective. They also need to consider their attitudes and attributions, and numerous other skill and competency areas. Thus, expanding beyond the 10 domains of school psychological practice discussed earlier, effective school psychologists also demonstrate many of the following attitudes and attributions:

- Accept responsibility for all children and adolescents in their schools
- Focus on the "whole" or entire student, looking at both strengths and weaknesses
- Deliver services based on student need and outcome, not on student label, disability, or diagnosis
- Identify the progress and successes accomplished in all initiatives, while seeking to expand and strengthen those successes
- Are committed to continuing education, professional development, and skill expansion
- Look at all challenges as learning and continued professional development opportunities

Effective school psychologists also demonstrate the following skills and competencies:

- Focus on comprehensive services at the systems, team, small group, and individual levels
- Recognize the psychological and education impact of gender, race, culture, socioeconomic status, and sexual orientation
- Recognize the psychological and educational impact of poverty, divorce, abuse and neglect, and other life crises
- Understand the characteristics of effective instructional environments and consult with staff to ensure that all students consistently experience such environments during their school careers
- Have good problem-solving skills with which they link functional assessment to strategic intervention
- Use acceptable assessment and evaluation procedures
- Can develop, implement, and evaluate effective behavioral and curricular interventions and programs
- Are guided by and effectively use data
- Are dedicated to using evidence- and research-based procedures and strategies

In summary, effective schools, effective school psychological services, and effective school psychologists are crucial in order to successfully impact all students. Guided by theory, research, application, and practice, psychology has and will continue to have a major impact on schools at the organizational, small group, and individual student and staff levels. This section of the encyclopedia reviews some of the most important and relevant areas of applied psychology as practiced by school psychologists. Critically, many of these areas are not unique to school psychology; they are similarly important to other specialties in applied psychology. This demonstrates the breadth and depth of psychology as a whole. As the science of human behavior, its many applications present myriad opportunities to positively affect people's lives.

4. THIS SECTION OF THE ENCYCLOPEDIA

The articles in this section of the encyclopedia are broadly representative of the many important topics relevant to school psychology. The articles, nonetheless, certainly are not exhaustive. The articles prepared can be clustered logically into six broad areas of practice or application.

The first cluster of entries focuses on the importance of looking at schools within the context of the community. Effective schools recognize that they are a microcosm of their communities and that they are strongest when building partnerships that embrace and involve community constituencies. Halsell Miranda addresses the importance of dealing with

children and families from diverse cultures, a critical issue given our heterogeneous society and the need for cultural sensitivity and cultural equity as applied to both school and community. Cowan extends this discussion by emphasizing the need to bring home–school collaboration to a conscious level.

Parents are their children's first teachers, and their support of the school process significantly impacts students' academic achievement. Parent involvement must be nurtured by our schools. But parent involvement does not necessarily mean that parents must be physically involved in school activities on school grounds. Parent involvement, at its core, means that parents support the educational process of their children and prepare them for school. In fact, if parents only did the following five things with their children, academic achievement would significantly improve across our country: promote their children's health and nutrition, read to their children and support their involvement in literacy activities at home, talk to their children so that their vocabulary (and brains) can develop and so that they feel emotionally supported, decrease their children's use of the television along with other media-based technologies, and support the school's instructional, including homework, processes.

Finally, Cowan also talks about school–community partnerships. Many times, our schools are not collaborating in effective and efficient ways with community resources. Collaboration with health and mental health agencies can increase students' health and readiness for school. Collaboration with government and public safety officials can increase school safety. And collaboration with volunteer agencies can bring additional resources into our schools. School–community partnerships reinforce for students that many individuals and groups are committed to and support their education. All of this helps students to recognize that they are not alone in this process, and that others are ready to celebrate their successes.

The second cluster focuses on the place of assessment and student evaluation in effective schools. Clearly, students need to be evaluated on an ongoing basis to ensure that they are mastering the academic goals and objectives embedded in the curriculum. As students' ability to management their own behavior increases their academic engagement, which then increases their potential for academic achievement, student evaluation must include behavioral evaluation as well as academic evaluation. Student evaluation becomes more focused on student assessment when students are not mastering their skills or when their behavior is interfering with academic progress. At this point, assessments should use functional assessment approaches to determine why the difficulty exists and, strategically, what interventions will most effectively remediate the situation. As reinforced by Ortiz, all assessment should be nondiscriminatory assessment. That is, all assessments should be devoid of bias (e.g., gender, racial, cultural, language, and geographic bias) and administered in objective and standardized ways. Nondiscriminatory assessment is essential to the validity of all assessment results. Without valid results, the potential for effective intervention will be compromised.

The other assessment areas that are investigated to identify the source of a student's difficulties reflect the specific outcomes that are expected of students. In general, these focus on students' academic, social, emotional, behavioral, and, sometimes, vocational outcomes. Thus, assessments can be done in the areas of intelligence and cognition, academics and achievement, behavior and emotions, and vocational interest and readiness. In this section, Hintze discusses behavioral assessment in schools; Evan, Williams, and Schultz discuss behavioral observation in schools; Miller discusses neuropsychological assessment in schools; and Levinson discusses vocational assessment in schools.

The third cluster focuses on the importance of characteristics of students with special needs. As noted, some students do not acquire academic and/or behavioral skills and require specific or strategic interventions. Some of these students have specific disabilities that can be accommodated for in their classrooms through minor curricular adaptations or specific supportive approaches (e.g., preferred seating, allowing oral instead of written responses, additional time for tests, lecture notes written by the teacher, tape-recorded books). Other students have specific disabilities that interfere with their school success so significantly that they are identified as special education students under the federal Individuals with Disabilities Education Act (IDEA) and receive specific interventions through an individualized education program (IEP). Under IDEA, 13 different disability areas are specified, including those that cover learning, behavior, cognition, health impairments, speech and language, hearing, vision, autism, and traumatic brain injuries. Some states also include students who are gifted and talented here, recognizing that these students do not have disabilities, but they do have exceptional needs. Additionally, some specific problem areas, such as attention deficit

disorders, may qualify for classroom accommodations and, if very serious, as a disability under the "other health impaired" category.

Regardless, effective schools are aware of the possible disability areas that interfere with a student's progress, how to assess them, how to respond to them through interventions, and how to include those students, to the greatest degree possible, in general education classrooms and in the general education curriculum. In this section, a number of these disabilities and related areas are presented: learning disabilities by Tilly; students with emotional and behavioral problems by Huberty; attention deficit disorders—school-based interventions by DuPaul and Tresco; and the teaching of gifted students by Feldhusen.

The fourth cluster focuses on the importance of academic interventions. Under the federal No Child Left Behind legislation, students' academic progress and achievement is the primary goal of the effective school. In order to meet this goal, teachers must be trained to be effective instructional coaches, the curriculum must be well matched to student styles and desired outcomes, and students must be fully engaged while being taught at their instructional levels. Academic success occurs because of effective planning and sound pedagogical instruction for all students. Academic success, however, also comes because specific interventions are implemented for those students who need more attention, practice, and alternative approaches to learning.

While effective schools focus on well-defined areas of instruction (e.g., reading and literacy, mathematics, science, social studies), they also focus on teaching students the skills they need to be independent thinkers and learners. Thus, metacognitive instruction is needed. This involves teaching students the "learning how to learn" process so that they can apply this process to novel learning situations. Metacognitive training can have a greater impact on students than the skill-based lessons of the traditional curriculum. As the proverb goes, it is often better to teach a person to fish than to simply give him or her fish on a daily basis.

In this section, Rafoth discusses the prevention of academic failure, emphasizing the proactive need to make all students successful, rather than waiting for students to fail and then reacting to their failure. Rathvon outlines a number of areas of academic interventions that should be considered when students are not mastering skills in timely ways. Joseph focuses more specifically on reading interventions, a major area of concern, especially during students' elementary school years. Finally, Gettinger addresses the importance of study skills instruction, part of the metacognitive foundation that helps students to become independent learners.

The fifth cluster focuses on the importance of discipline, behavior management, and school safety. Clearly, if students do not have the behavioral skills that support the academic process (e.g., listening, following directions, working cooperatively in groups, demonstrating task persistence), they will have more difficulty with academic engagement and achievement. Similarly, if students do not have good interpersonal and conflict resolution skills, and if their schools are unsafe and unsettled, their necessary concern with resolving their conflicts and protecting their own welfare will take precedence over school work. Thus, positive, safe, and predictable school environments are an essential component of effective schools.

To provide guidance in this important area, Bear talks about the component parts of an effective school discipline and behavior management approach. Larson addresses this topic in a different way, focusing on school violence prevention. Furlong and his colleagues then investigate a more specialized situation, bullying and abuse on school campuses, that schools must respond to both legally and ethically. Finally, Elliott and Lang discuss the importance of social skills training as the anchor to a positive and safe school. Expanding briefly, just as students need to learn skills to be academically successful, so too do they need to learn prosocial skills in order to be behaviorally and interpersonally successful.

The sixth, and final, cluster focuses on the importance of health and mental health prevention and intervention. Once again, when students are not physically and mentally healthy, their school success potentially suffers. In this section, Copeland and Crepeau-Hobson discuss the importance of health promotion in the schools. Although instruction in this area often occurs in many schools, the lessons still must be actively reinforced by students, parents, and staff in order for them to be consistently practiced. Poland then addresses suicide intervention and schools. Unfortunately, the United States has a rate of student suicide that is higher than that of any other industrialized country. This topic, then, when combined with the health promotion article, will hopefully alert psychologists and educators to the importance of attending to students' health and mental health needs and to the action steps that are needed, in the same context, to prevent student suicide.

5. CONCLUSION

School psychology dates back to the early 1920s relative to training and practice, and to the mid-1950s at a professional organizational level. While integrating research and clinical practice from a number of other psychology specialty areas, as well as from education, school psychology has its own research base that it can distinctly call its own. Similarly, while other psychology specialty areas do indeed contribute to the application of psychology in the schools, school psychologists have a primary professional responsibility to understand the culture, norms, and contexts of education so that the best interventions and services are used. School psychologists can best serve their schools, staffs, and students when they are problem solvers, consultants, and intervention specialists. In the end, even in the face of so many needs and so many challenges, school psychologists contribute both to the science and practice of effective schools and successful students.

See Also the Following Articles

Achievement Tests ■ Behavioral Assessment in Schools ■ Behavioral Observation in Schools ■ Consultation Processes in Schools ■ Diverse Cultures, Dealing with Children and Families from ■ Educational Achievement and Culture ■ Educational and Child Assessment ■ Effective Classroom Instruction ■ Exceptional Students ■ Full-Service Schools ■ Gender and Education ■ Gifted Students ■ Health Promotion in Schools ■ Home–School Collaboration ■ Intelligence Assessment ■ Learning ■ Learning Disabilities ■ Learning Styles and Approaches to Studying ■ Nondiscriminatory Assessment in Schools ■ Special Education ■ Teaching Effectiveness ■ Transfer of Learning

Further Reading

Arkansas Department of Education–Special Education. (2003). *A portfolio approach to conducting focused monitoring in selected Arkansas school districts*. Little Rock, AR: Arkansas Department of Education.

Berliner, D. C. (1988). Effective classroom management and instruction: A knowledge base for consultation. In J. L. Graden, J. E. Zins, & M. J. Curtis (Eds.), *Alternative educational delivery systems: Enhancing instructional options for all students* (pp. 309–326). Silver Spring, MD: National Association of School Psychologists.

Bickel, W. E. (1999). The implications of the effective schools literature for school restructuring. In Reynolds, C. R., & Gutkin, T. B. (Eds.), *The handbook of school psychology* (3rd ed., pp. 959–983). New York: John Wiley & Sons, Inc.

Cawalti, G. (Ed.) (1995). *Handbook of research on improving student achievement*. Arlington, VA: Educational Research Service.

Chall, J. S. (2000). *The academic achievement challenge: What really works in the classroom?* New York: Guilford Press.

Gettinger, M., & Stoiber, K. C. (1999). Excellence in teaching: Review of instructional and environmental variables. In Reynolds, C. R., & Gutkin, T. B. (Eds.), *The handbook of school psychology* (3rd ed., pp. 933–958). New York: John Wiley & Sons, Inc.

Knoff, H. M. (2004). Inside Project ACHIEVE: A comprehensive, research-proven whole school improvement process focused on student academic and behavioral outcomes. In K. Robinson (Ed.), *Advances in school-based mental health: Best practices and program models*. Kingston, NJ: Civic Research Institute, Inc.

Knoff, H. M. (2000). Organizational development and strategic planning for the millennium: A blueprint toward effective school discipline, school safety, and crisis prevention. *Psychology in the Schools*, *37*, 17–32.

Knoff, H. M., Hines, C. V., & Kromrey, J. D. (1995). Finalizing the Consultant Effectiveness Scale: An analysis and validation of the characteristics of effective consultants. *School Psychology Review*, *24*, 480–496.

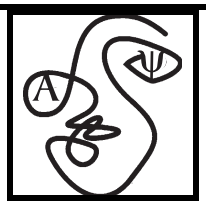
Wang, M. C., Haertel, G. D., & Walberg, H. J. (1990). What influences learning? A content analysis of review literature. *Journal of Educational Research*, *84*, 30–43.

Wang, M. C., Haertel, G. D., & Walberg, H. J. (1993/1994). What helps students learn. *Educational Leadership,* December/January, 74–79.

Ysseldyke, J., Dawson, P., Lehr, C., Reschly, D., Reynolds, M., & Telzrow, C. (1997). *School psychology: A blueprint for training and practice II*. Bethesda, MD: National Association of School Psychologists.

Ysseldyke, J., & Elliott, J. (1999). Effective instructional practices: Implications for assessing educational environments. In Reynolds, C. R., & Gutkin, T. B. (Eds.), *The handbook of school psychology* (3rd ed., pp. 497–518). New York: John Wiley & Sons, Inc.

Ysseldyke, J. E., & Christenson, S. L. (2002). *Functional assessment of academic behavior: Creating successful learning environments*. Longmont, CO: Sopris West.



School-to-Work Transition

David L. Blustein, Justin C. Perry, and David B. DeWine

Boston College, Chestnut Hill, Massachusetts, USA

1. Overview
2. A Historical and Sociopolitical Perspective
3. The Psychological Realm
4. Conclusion
   Further Reading

GLOSSARY

connecting activities Psychological and educational interventions designed to foster the connection between school and work.

school-based learning Educational experiences rooted in the classroom or school setting.

school-to-work Focusing primarily, but not exclusively, on non-college-bound youth, school-to-work (STW) generally refers to the transition from high school into the workforce and/or post-secondary training. This area of inquiry, which has explored the economic, moral, sociological, and psychological factors associated the STW process, has yielded important research findings along with suggestions for the improvement of educational and psychological practice.

work-based learning Educational experiences rooted in an occupational setting.

work-bound youth Those individuals who do not attend a 4-year college immediately after graduating from high school.

This article summarizes the school-to-work (STW) transition movement with respect to its social and economic origins, current knowledge base, and implications for research and practice. The authors first provide a historical overview of the STW movement and describe its underpinnings in vocational education and career education. The relationship between the STW transition and psychological theory and research is then discussed. Conclusions are given for scholars and practitioners in vocational psychology.

1. OVERVIEW

Broadly speaking, the focus on the school-to-work (STW) transition, which dates to the late 1980s, refers to a movement of educational, social, political, and public policy reform that is designed to help students transition from high school to the world of work or post-secondary training. The initial emphasis of this movement was directed toward the needs of students who do not graduate from 4-year colleges. However, as the movement evolved in the United States, it was expanded to include youth who were transitioning to college. This transition is a highly challenging and critical task for adolescents, with significant relevance to both academic and nonacademic developmental trajectories. The STW movement has attracted much interest in several disciplines, most notably life course sociology, counseling psychology, and life span developmental psychology.

Compared to other countries throughout the world, the overall lack of a systematic approach to the STW transition in the United States has resulted in its diminished capacity to prepare work-bound youth for entry into the labor market. Youth in the United States,

©2004 Elsevier Inc.
All rights reserved.

particularly those who do not attend a 4-year college, are typically left to navigate a period of "floundering" in which individual agency and the availability of resources play a major role in determining the level of risk and potential for an adaptive transition. In comparison, many European countries, such as Finland and Germany, have highly formal, unified, and institutionalized STW systems. Although these STW initiatives have a number of assets, including the reduction of floundering after high school, some critics have noted that these highly structured transition processes risk reinforcing social and economic inequality.

Current programmatic and research efforts on the STW transition stem from a number of interrelated problems. The most salient sources of impetus for STW initiatives were (i) the emergence of declining wages, beginning in the early 1970s, among skilled and unskilled workers; (ii) the increasing level of automation and computer technology, reducing the need for semiskilled and unskilled workers; (iii) the growing disparities in wealth between college-educated and non-college-educated individuals; and (iv) the increasing debates about the quality, relevance, and standards of education. These issues eventually led to awareness by policy analysts, educators, and social scientists that the U.S. economy is experiencing a labor crisis of low skills and high wages. Jobs with reasonably decent wages that were traditionally available for high school graduates in industrial, manufacturing, and agricultural fields are no longer realistic opportunities for work-bound youth due to lack of availability and security. For adolescents from inner cities and rural areas, this crisis has become especially pronounced; for instance, the unemployment rate for urban youth of color is approximately double that of their European American peers attending schools in suburban areas. Nonetheless, numerous STW policy statements during the past decade have been directed toward students from all socioeconomic backgrounds and school systems, including the college-bound.

2. A HISTORICAL AND SOCIOPOLITICAL PERSPECTIVE

The historical and social antecedents of the STW movement can be traced to various educational and policy-based ideas and programs during the past several decades. Soon after the turn of the 20th century, the classical approach to secondary education that focused on the core areas of language, math, science, and history (which had been available primarily to the affluent) began to be replaced by a philosophy that sought to train an increasing proportion of students for employment roles, largely in industry and agriculture. This shift in the purpose of U.S. education gave rise to the vocational education movement and later to career education. Such a fundamental change presaged the current debate pertaining to the objectives and pedagogical approaches of today's school systems.

In seeking to provide high school students with practical, hands-on skills that could lead to direct employment, the vocational education movement divided academic training along a vocational education versus college-preparation track. This bifurcation occurred mainly because of the changing demands of a mass-producing and consumer-driven society, in which employers increasingly needed workers with technical skills who could perform tasks efficiently. As a result, vocational education was often viewed as a second-rate educational alternative designed to meet the demands of industries, covertly perpetuating class boundaries. According to many educators and policy analysts, however, vocational education has not been a successful system. In order to rectify the dualistic systems of learning established by this movement, career education emerged as a rubric for a wide array of programs that share two interrelated objectives: to enhance the occupational relevance of education and to reduce the social class distinctions between vocational and academic tracks. The major mechanism of career education, therefore, was the infusion of a career development and occupational focus into all levels of the traditional academic curriculum.

In the early 1970s, career education received considerable public attention and funding. Although much of the financial resources and intellectual energy within this movement have dissipated, many of the ideas were integrated into the schools and have formed part of the conceptual framework for the STW movement. Specifically, the discussion of the vast economic and moral implications of existing divisions between college-bound and non-college-bound, a dialogue initially set in motion by John Dewey in the early 1900s, was clearly articulated within the career education movement. This discussion helped set the stage for the debates that are currently taking place with regard to the best ways of reconciling the contradictory aims of secondary school education in the United States. Namely, this contradiction revolves around the notion of academic rigor versus academic relevance, or the disconnection between real-world applications of learning and the conventional "skills and drills" taught in math, English, and science.

3. THE PSYCHOLOGICAL REALM

The literature in psychology that seeks to understand how the STW transition process occurs represents a critical body of knowledge that has the potential to inform subsequent theory and practice. A broad examination of the literature suggests that a wide array of definitions of an adaptive STW transition have been used. In general, social scientists have tended to rely on global indices of employment and unemployment or level of occupational attainment achieved in the labor market. In contrast, psychological investigations have pointed to definitions encompassing internal or relatively proximal characteristics of one's posttransition behavior, including job satisfaction and occupational choice congruence, development of adaptive coping skills on the job, and development of self-concept. Based on a review of recent literature on the STW transition, an adaptive transition may be best understood as including some internal psychological index of the degree of individual satisfaction one experiences with respect to work and nonwork roles as well as more global and objective indices of how well the individual is faring in attaining access to opportunities for growth, advancement, and further training.

When considered collectively, the existing research has revealed that four broad factors are associated with an adaptive STW process:

- Competencies in basic academic skills and vocationally relevant domains
- Self-initiative, planfulness, flexibility, and purposefulness
- An active and supportive relational environment, including family, positive peers, teachers, counselors, and coworkers and supervisors
- An educational environment that offers a clear connection to the world of work

Post hoc observational studies have revealed several key psychological attributes associated with successful STW transitions that represent these characteristics. Furthermore, studies have found that families who exhibit characteristics such as cohesiveness, expressiveness, and democratic decision making tend to foster important career-oriented learning processes for adolescents.

Researchers have also shown that the sequence of occupational activities in which a young adult engages can affect his or her psychological well-being. For example, workers from European samples who move from employment to additional training exhibit higher levels of psychological well-being, whereas young adults who move from school into the military, with no prior employment history, experience lower levels of psychological well-being.

Finally, longitudinal research data indicate that young adults who experience increased floundering are at risk for higher levels of psychological stress. Floundering refers to employment instability in which a young adult or adolescent changes jobs with little or no support or direction. Although a certain amount of occupational exploration can influence an adaptive transition by increasing person–environment awareness, too much instability has been linked to negative psychosocial outcomes.

As an integral aspect of gaining a richer understanding of transitioning from school to work, social scientists also have sought to understand how membership in a marginalized group might influence the STW process. Specifically, scholars have investigated how youth with disabling conditions, young women, and youth of color negotiate the complex transition in the context of less access to the educational and economic opportunity structure. Despite the sparseness of research in these areas, initial studies have suggested that disabling conditions, gender, and visible racial and ethnic minority status play a discernible and aversive, although not uniform, role in the STW transition. That is, the available evidence indicates that the transition from school to work offers a more pronounced set of challenges and barriers for these minority groups compared to members of the dominant culture.

4. CONCLUSION

One of the primary implications of the STW movement within applied psychology has been its explicit focus on the needs of poor and working-class students. Indeed, discourse and research initiatives stimulated by the STW movement have placed poor and working-class students at the forefront of thinking in vocational psychology. Moreover, the STW movement has highlighted the need for an active consideration of social class, race, and other social barriers in constructing educational policy, building effective theories, and designing research.

Furthermore, research concerning the STW transition has informed counseling practioners of the need to examine their own implicit class biases in how they conceptualize work and in how they design interventions. Traditional office-based counseling practice may no longer be an effective way to engage work-bound youth in the process of negotiating adaptive transitions

to work and post-secondary training. The STW movement has highlighted the need for scholars and practitioners in vocational psychology to develop a broader understanding of issues pertaining to equity, structural aspects of the opportunity structure, and the social context of work, as well as the ability to communicate with larger systems, such as school and industry.

See Also the Following Articles

Holland's Theory (Vocational Personality Types) ■ Indecision, Vocational ■ Vocational Assessment in Schools ■ Vocational Interests ■ Vocational Psychology, Overview ■ Work Adjustment

Further Reading

Blustein, D. L., Juntunen, C. L., & Worthington, R. L. (2000). The school-to-work transition: Adjustment challenges of the forgotten half. In S. D. Brown, & R. W. Lent (Eds.), *Handbook of counseling psychology* (3rd ed., pp. 435–470). New York: Wiley.

Charner, I., Fraser, B. S., Hubbard, S., Roger, A., & Horne, R. (1995). Reforms of the school-to-work transition: Findings, implications, and challenges. *Phi Delta Kappan*, *77*, 58–59.

Gray, K. C., & Herr, E. L. (1998). *Workforce education: The basics*. Needham Heights, MA: Allyn & Bacon.

Marshall, R., & Tucker, M. (1992). *Thinking for a living: Education and the wealth of nations*. New York: Basic Books.

Olson, L. (1997). *The school-to-work revolution*. Reading, MA: Addison-Wesley.

O'Neil, H. F. (1997). *Workforce readiness: Competencies and assessment*. Mahwah, NJ: Erlbaum.

Pautler, A. J. (Ed.). (1994). *High school to employment transition: Contemporary issues*. Ann Arbor, MI: Prakken.

Pauly, E., Kopp, H., & Haimson, J. (1995). *Homegrown lessons: Innovative programs linking school and work*. San Francisco: Jossey-Bass.

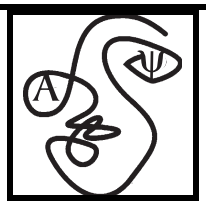
Resnick, L. B., & Wirt, J. B. (1996). *Linking school and work: Roles for standards and assessments*. San Francisco: Jossey-Bass.

Savickas, M. L. (1999). The transition from school to work: A developmental perspective. *Career Development Quarterly*, *47*, 326–336.

Secretary's Commission on Achieving Necessary Skills. (1992). *Learning a living: A blueprint for high performance*. Washington, DC: U.S. Department of Labor.

Wilson, W. J. (1996). *When work disappears: The world of the new urban poor*. New York: Random House.

Worthington, R. L., & Juntunen, C. L. (1997). The vocational development of non-college-bound youth: Counseling psychology and the school-to-work transition movement. *The Counseling Psychologist*, *25*, 323–363.



School Violence Prevention

Jim Larson
University of Wisconsin, Whitewater, Wisconsin, USA

1. Introduction
2. Definition of School Violence
3. Historical Perspective
4. Effects of School Violence
5. Organizing Structures for School Violence Prevention
6. An Integrated Approach to School Violence Prevention
7. Conclusion
Further Reading

GLOSSARY

bullying Continual infliction of harm or threat of harm, physical or psychological, within the context of a clear imbalance of power.

hostile excluding A form of nonphysical aggression, more common in females, that involves efforts to cause psychological distress through directed exclusion from a social group or desirable social activities.

protective factor Characteristics of the individual or environment that serve to moderate the impact or improve people's resistance to human dysfunction.

risk factor Characteristics of the individual or the environment that serve to increase the probability of onset, degree of severity, or duration of human dysfunction.

zero tolerance A school discipline policy intended primarily as a method of sending a message that certain violence-related behaviors will not be tolerated by punishing both major and minor offenses severely.

School violence prevention involves the application of universal, selected, and indicated evidenced-based programs and procedures in the school setting in order to reduce the incidence of interpersonal violence in all its forms.

1. INTRODUCTION

Efforts to prevent school shootings such as those that occurred at Columbine High School and elsewhere are only one facet of a much larger continuum of school violence prevention. The statistical probability of being shot by a fellow student is extraordinarily remote, but the day-to-day likelihood of being threatened by a bully, punched by a classmate, or victimized by mean-spirited teasing is not. Understanding and appreciating the myriad contextual, intrapersonal, developmental, and educative aspects of maintaining a safe and supportive learning climate for both students and adults is the focus of this article.

2. DEFINITION OF SCHOOL VIOLENCE

School violence involves all the behaviors along the full continuum of interpersonal physical, verbal, and psychological aggression that interfere with the rights

©2004 Elsevier Inc.
All rights reserved.

of all children to learn and the rights of all adults to work in a safe educational environment. School violence is much more than just the rare headline-grabbing multiple homicides; it involves the quality of the day-to-day interpersonal relationships in the context of the school setting—relationships that contribute to or detract from everyone's sense of personal safety and well-being. Teasing, bullying, threatening, hostile excluding, drug dealing, and fighting are example behaviors of concern. This definition also recognizes that the continua of physical, verbal, and psychological aggression have intensity, frequency, and emotional and physical injury plot points. For example, mean-spirited teasing or bullying directed at a target student can be examined with regard to the intensity of the interaction, the frequency of occurrence, and the emotional and physical consequences to the victim.

Additionally, it is important to understand that school violence may also be viewed from a skills-deficit perspective. A child who is resorting to physical aggression to resolve a concern may be demonstrating a lack of skill in social problem solving and anger management. From this viewpoint, school violence can also be defined as an index measuring the deficiency of prosocial skills and nonviolent conflict resolution strategies among individuals in the context of a school setting. Framing the definition in this fashion has useful implications for prevention and intervention policies and procedures.

3. HISTORICAL PERSPECTIVE

For many educators responsible for student safety before the 1970s, preventing school violence meant little more than administratively ridding the school of the most potentially violent students. When efforts to intervene failed, children and youth with chronic aggressive anger problems, gang affiliations, or other problematic characteristics were summarily dismissed from school. Subsequent federal legislation, such as Public Law 94-142 and the Education of All Handicapped Children Act of 1975 (now Public Law 105-17, the Individuals with Disabilities in Education Act Amendments of 1997), and Supreme Court decisions such as *Goss v. Lopez* changed the way students with challenging behaviors needed to be addressed. The 14th Amendment to the Constitution holds that no state shall "drive any person of life, liberty, or property, without due process of law." Ruling in the 1975 *Goss v. Lopez* case, the Supreme Court held that to deprive an individual of an education is to deprive that individual of a property right. Barring or otherwise excluding any individual from a free and appropriate public education without due process was determined to be unconstitutional. The effect of Public Law 94-142 and the Goss decision was to place the obligation on the schools to address the needs of everyone who desired a public education, including those with challenging behaviors. The days of simply excluding aggressive students without due process were over.

In 1978, the federal government published the first large-scale report on school safety, *Violent Schools–Safe Schools: The Safe School Study Report to Congress*, and the nation's eyes were opened to a new and disquieting feature of public education. Among the findings were that in a given month at school, approximately 282,000 (1.3%) students were physically attacked in secondary schools, almost 8% of urban junior and senior high school students missed at least 1 day of classes because they were afraid to go to school, and nearly 5200 of the nation's million secondary school teachers were physically attacked. For most middle-class Americans, however, the issues of violence in school remained pretty much an urban concern: The Blackboard Jungle was a long way from the comparative safety of suburbia. The major urban school districts spent the 1970s and 1980s fighting the escalating street crime that spilled into their hallways using every method at their disposal, including locker searches, armed guards, and metal detectors. In 1984, a presidential directive established the National School Safety Center to serve as an information clearinghouse and resource center for beleaguered schools. In 1986, Boston public health physician Deborah Prothrow-Stith introduced the nation's first dedicated violence prevention curriculum, Violence Prevention Curriculum for Adolescents, into the Boston school system. In 1987, the Seattle-based company Committee for Children followed suit with its program for younger students, Second Step: A Violence Prevention Curriculum. In 1988, the Milwaukee Public Schools initiated the Violence Reduction Program, the nation's first districtwide, multisystemic effort to address the mounting problem of student violence through curricular, counseling, and parent training efforts. In suburban, small town, and rural school districts outside of the larger cities, however, schools had not yet even begun to lock their doors.

The relative complacency of the nonurban U.S. public school was changed forever on January 17, 1989,

when Patrick Purdy, 26, armed with an AK-47 assault rifle, opened fire on a playground at Cleveland Elementary School in Stockton, California. Five children died and 29 children and one teacher were wounded before Purdy killed himself. Whereas this assailant was an adult, it was the homicides perpetrated in the school setting by students in communities such as Jonesboro, Arkansas, Springfield, Oregon, and Littleton, Colorado, that provided the greatest impetus to the development of school violence prevention programs. Since the early 1990s, concerns both real and imagined about school shootings have fueled both research and product development in the area of school violence prevention.

Somewhat lost in all of the media coverage of these high-profile homicides were the historical facts regarding school violence. National surveys by the Bureau of Justice Statistics and the National Center for Educational Statistics indicated that there was a 46% decline in the nonfatal violent victimization rate (simple assault and more violent crimes, including rape, robbery, and aggravated assault) in schools between 1992 and 2000 among students ages 12–18. For all categories of crime victimization except theft, young people were safer at school than away from it. The number of teachers who were threatened with violence has also declined during the past decade, although the number of teachers who were actually assaulted has remained steady at approximately 4%. Data from the National School Safety Center's School-Associated Violent Deaths database showed that the incidence of single-victim homicide in schools also steadily declined during the past decade. Indeed, during the 1998–1999 school year that included the Columbine shootings, U.S. schools had the third lowest total of school-related homicides since 1992–1993. The statistic that increased during that same period was multiple homicide events in the school setting (Table I).

Although the number of homicides and serious criminal victimization in schools continues to decline, the same cannot be said for other forms of school violence. For instance, student reports of bullying are on the increase. It is estimated that during their school careers, one in three students will be involved in school bullying incidents as a victim, perpetrator, or both. Conversely, the odds of being killed in a school shooting have been estimated to be approximately 1 in 2 million, or approximately the same as being struck by lightning. This has implications for designers of violence prevention programs, who must account for base rates along the continuum of school violence if the effort is to be meaningful in the lives of most students. Planning effectively for violence prevention takes into account that time and money spent preparing for a monumentally unlikely event at the expense of real day-to-day violence and victimization is both scientifically and fiscally indefensible.

TABLE I
School Deaths Attributable to Single and Multiple Homicides

School year	*Deaths by single homicide*	*Deaths by multiple homicide*
1992–1993	42	5
1993–1994	42	4
1994–1995	17	0
1995–1996	20	10
1996–1997	21	4
1997–1998	18	17
1998–1999	9	17
1999–2000	16	6
2000–2001	16	2
2001–2002	2	0

4. EFFECTS OF SCHOOL VIOLENCE

How can a student be expected to concentrate on academic learning when the potential for physical assault or the next round of mean-spirited teasing is just around the corner? What is it like to dread recess or suppress the urge to drink water for fear of the need to later have to venture into a dangerous restroom? How can a teacher possibly provide his or her full measure of skill when there is good reason to be frightened of some of the students?

Violence in the school setting has multiple physical, psychological, and educational effects on individuals and has a deleterious effect on the learning climate for everyone. Sociobiological theories of motivation posit that humans have four basic needs: drives to acquire, bond, learn, and defend themselves from harm. Students and teachers who must spend time and cognitive energy on concerns for their own safety cannot simultaneously devote their motivation and attention to the educational process in the most productive manner. Safe school environments allow students the freedom to attend to learning tasks with greater creativity, academic risk taking, and peer affiliation. Research from the Effective

Schools movement of the 1970s and, recently, from Project ACHIEVE has empirically demonstrated the positive effects on academic climate and student achievement through broad-scale measures to create safe, disciplined learning environments. Project ACHIEVE is a nationally known and evidence-based school reform and improvement program that has been implemented in more than 1500 school or districts throughout the United States. It has been designated as a Model Prevention Program by the Office of Juvenile Justice and Delinquency Prevention in the U.S. Department of Justice; the Collaborative for Academic, Social, Emotional Learning; and the Substance Abuse and Mental Health Services Administration of the Department of Health and Human Services.

5. ORGANIZING STRUCTURES FOR SCHOOL VIOLENCE PREVENTION

Effective school violence prevention programs are multisystemic endeavors that are integrated into the fabric of the educational process, not simply added on as time or budget permit. Planning and implementing procedures for the safety and emotional well-being of students are no less thoughtfully or rigorously studied than any other critical aspect of the educational process. Prevention research performed during the past three decades has emphasized the development of stakeholder partnerships. Individual classrooms exist in buildings, which exist in school systems located in communities. Consequently, the violence prevention process should have an identifiable presence and linkage across community, school district, and school building levels of policy and organization. Table II provides some of the basic roles and functions in school violence prevention at these levels.

TABLE II
Basic Levels, Roles, and Functions in School Violence Prevention

Community level
Address community risk factors associated with youth violence
Enact and enforce juvenile criminal codes
Enact and enforce strong antidrug and firearm laws
Establish and coordinate community emergency response procedures
Provide emergency training expertise to school staff
Provide expertise and coordinate school security assessments
Provide safe after-school activities for all students
Provide expertise in research-supported prevention measures
School district level
Provide school board mandate for safe schools
Secure funding through budgetary process
Facilitate coordination with community resources
Ensure appropriate school mental health services
Develop policies for school security and emergency response
Facilitate staff training in:
Classroom discipline and problem-solving methods
Antibullying procedures
Instructional design and curricular modification
Emergency/security procedures
School building level
Establish representative school safety planning team
Conduct broadly based needs assessments
Establish primary, secondary, and tertiary prevention plans
Enact collaboratively developed school discipline plan
Practice emergency response procedures
Facilitate ongoing staff development

6. AN INTEGRATED APPROACH TO SCHOOL VIOLENCE PREVENTION

School violence prevention measures aim to (i) prevent the onset of violent behavior in the general school population, (ii) identify and intervene with those students who are at risk for violent behavior, and (iii) reduce the frequency or intensity of violence in those students who are currently engaging in violent behavior. Compared to more reactive approaches, such as "zero tolerance" with its sanctions-based emphasis on suspension and expulsion, a prevention approach is primarily skills based and education oriented. Prevention approaches do not completely eschew the use of exclusionary procedures, but such reactive measures are viewed more as a last line of response and typically as a protective rather than intervention measure. Prevention approaches draw from numerous literature bases, including those from risk and protective factors, social context behavior, child development, and learning theory, to create a school

ecology conducive to healthy interactions and nonviolent conflict resolution.

Because students learn and interact in several social contexts and with numerous individuals, these procedures are often multiply targeted. For instance, a child with anger control problems may receive anger management skills training, and simultaneously his or her parents may be engaged in parent training while his or her teacher is learning more effective classroom discipline strategies. Targets for school violence prevention include the student, teacher, family, classroom, school, school district, neighborhood, community, state, and national policy.

Fully integrated prevention programs target risk factors and skill development in multiple contexts and among multiple populations. In doing so, programs use universal, selected, and indicated prevention procedures.

6.1. Universal Prevention Measures

When school violence prevention targets an entire population, such as a whole classroom or school, the endeavor is known as primary prevention and the measures used are referred to as universal prevention measures. Universal prevention measures are designed for all the individuals in the targeted population—students, teachers, or parents—regardless of each individual's risk status. At the classroom level, a teacher may provide the students with curricular instruction on the subjects of conflict resolution strategies, ways to manage anger, or safe handling of firearms. Because of its broad reach, universal classroom prevention curricula place primary emphasis on teaching the students about risk and prevention at a basic knowledge level rather than actually training them in complex behavioral skills. At the preschool to early primary levels, students may acquire problem-solving language and learn to name and identify feeling states in themselves and others. In the later elementary years, students may be taught to understand procedures for social problem solving, conflict resolution, and anger control.

With the whole school as a target for primary prevention, a school administration may offer in-service workshops to its teaching staff to bring them up to date on the latest techniques in classroom discipline skills or to help them learn to diversify instruction for behaviorally challenging students. A universal prevention measure that targets use of the physical environment may involve brightening the common areas with new décor and additional lighting, modifying class changing times to reduce hallway crowding, or locking entranceways after the start of school.

6.2. Selected Prevention Measures

Within school populations are individuals or groups that, because of personal characteristics or environmental factors, have elevated risk. Efforts to meet the needs of these individuals are known as secondary prevention, and the procedures used are referred to as selected prevention measures. One of the goals of secondary prevention in the school setting is to identify those individuals within a population who may be engaging in precursor behaviors that have predictive value for the onset of later, more serious concerns. Physically aggressive or bullying behavior, oppositional behavior, lying, or stealing, when observed in kindergarten or early primary-level children, are known to have a measure of predictive validity for later, more serious violent or antisocial behavior. In addition to receiving universal prevention measures, these children select themselves out for more intensive skill development.

Central to the successful application of selective prevention measures is the valid and reliable identification of students who will most likely benefit from the intervention. This is particularly important in situations in which there is significant environmentally based risk (e.g., a large population of low socioeconomic status children) and limited human resources to facilitate intervention. Multiple-gating methods allow classroom teachers and parents to use instruments such as risk-behavior checklists to narrow the population to those students or groups of students most likely to benefit from a particular selected measure.

Once identified, students receiving selected prevention measures focus on behavioral skill development, typically in an individual or small group setting under the direction of a school psychologist or guidance counselor. In these sessions, students learn to apply behavioral skills such as social problem solving, anger control, consequential thinking, or nonviolent conflict resolution to simulated scenarios in the training setting. In contrast to universal prevention, selected prevention measures go beyond merely knowing about violence prevention to actually learning and practicing critical behavioral skills, both in and out of the training setting.

Bullying prevention programs are examples of a selected measure applied at the whole school level. Longitudinal research indicates that students who

engage in bullying behavior in the elementary school are at high risk for continuing antisocial behavior in later life. In order to address the needs of both the bully and the victims of bullying, empirically supported bullying prevention programs train teachers, administrators, and students in role-appropriate identification, resistance, and effective intervention strategies.

6.3. Indicated Prevention Measures

The mission of the public schools includes providing an education to all students, regardless of their individual risk characteristics. Tertiary prevention services are designed to address the needs of students who display persistent, severely aggressive behaviors that have been unresponsive to universal or selected prevention procedures. These students are candidates for indicated prevention measures. Indicated measures are designed to reduce the frequency or severity of aggressive behaviors in the school setting—in effect, to keep the problem from worsening and protect the student and those around him or her. A tertiary prevention plan is typically formulated by a team of school and community professionals (e.g., school psychologists, mental health professionals, and agency representatives) working collaboratively with the student's parents. The student frequently receives a variety of services in and out of the school setting in a "wraparound" organizational structure. Multiple supports and services are "wrapped around" the student and, in many cases, the family in an effort to prevent problems from advancing to a more serious level. Indicated measures may involve skills training and individualized academic programming in the general or alternative school setting paired with family and community interventions.

7. CONCLUSION

Preventing school violence is a critical endeavor that requires training, organization, and ongoing commitment on the parts of those individuals responsible for its implementation. A fully integrated approach to school violence prevention brings together an understanding of childhood risk factors, skilled school and community professionals, and evidence-supported universal, selected, and indicted procedures to create a safe learning environment for students and staff members.

See Also the Following Articles

Aggression and Culture ■ Bullying and Abuse on School Campuses ■ Emotional and Behavioral Problems, Students with ■ School Discipline and Behavior Management

Further Reading

Aber, J. L., Brown, J. L., & Jones, S. M. (2003). Developmental trajectories toward violence in middle childhood: Course, demographics, differences, and response to school-based intervention. *Developmental Psychology*, *39*, 324–348.

Brock, S. E., Lazarus, P. J., & Jimerson, S. R. (Eds.). (2002). *Best practices in school violence prevention and intervention*. Silver Spring, MD: National Association of School Psychologists.

Brock, S. E., Sandoval, J., & Lewis, S. (2001). *Preparing for crises in the school: A manual for school crisis response teams* (2nd ed.). New York: Wiley.

Elliott, D. S., Hamburg, B. A., & Williams, K. R. (Eds.). (1998). *Violence in American schools: A new perspective*. New York: Cambridge University Press.

Hawkins, D. J., Herrenkohl, T., Farrington, D. P., Brewer, D., Catalano, R. F., & Harachi, T. W. (1998). A review of the predictors of youth violence. In R. Loeber, & D. P. Farington (Eds.), *Serious and violent juvenile offenders: Risk factors and successful interventions* (pp. 106–146). Thousand Oaks, CA: Sage.

Nohria, N., Lawrence, P., & Wilson, E. (2001). *Driven: How human nature shapes our choices*. San Francisco: Jossey-Bass.

Olweus, D. (1994). *Bullying at school: What we know and what we can do*. Malden, MA: Blackwell.

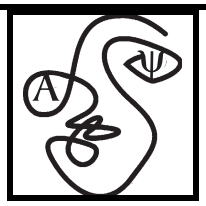
Pepler, D. J., & Slaby, R. G. (1994). Theoretical and developmental perspectives on youth and violence. In L. D. Eron, J. H. Gentry, & P. Schlegel (Eds.), *Reason to hope: A psychosocial perspective on violence & youth* (pp. 27–58). Washington, DC: American Psychological Association.

Shafii, M., & Shafii, S. L. (Eds.). (2001). *School violence: Assessment, management, prevention*. Washington, DC: American Psychiatric Publishing.

Shaw, D. S., Gilliom, M., Ingoldsby, E. M., & Nagin, D. S. (2003). Trajectories leading to school-age conduct problems. *Developmental Psychology*, *39*, 189–200.

Walker, H. M., & Epstein, M. H. (Eds.). (2001). *Making schools safer and violence free: Critical issues, solutions, and recommended practices*. Austin, TX: Pro-Ed.

Walker, H. M., Horner, R. H., Sugai, G., Bullis, M., Sprague, J. R., Bricker, D., & Kaufman, M. J. (1996). Integrated approaches to preventing antisocial behavior patterns in school-age children and youth. *Journal of Emotional and Behavioral Disorders*, *4*, 194–209.



Self-Confidence in Athletes

Robin S. Vealey
Miami University, Oxford, Ohio, USA

1. Definitions of Self-Confidence
2. Importance of Self-Confidence to Athletes
3. Sources of Self-Confidence in Athletes
4. Self-Confidence within Sport Teams
5. Measuring Self-Confidence in Athletes
6. Strategies to Enhance Self-Confidence in Athletes
Further Reading

GLOSSARY

home advantage Research finding that home teams win more than 50% of their competitions played under a balanced home-and-away schedule, often cited as a source of confidence for athletes.

optimism Tendency to expect the best possible outcome or dwell on the most hopeful aspects of a situation.

psychological momentum Athletes' perceptions that something has occurred that increases their probability of success, which typically creates a surge of confidence.

SC-cognitive efficiency An athlete's belief or degree of certainty that he or she can mentally focus, maintain concentration, and make effective decisions to perform successfully in sport.

SC-physical skills and training An athlete's belief or degree of certainty about his or her ability to execute the physical skills necessary to perform successfully in sport.

SC-resilience An athlete's belief or degree of certainty that he or she can regain focus after performance errors, bounce back from performing poorly, and overcome doubts, problems, and setbacks to perform successfully in sport.

self-efficacy Beliefs in one's capabilities to organize and execute the courses of action required to produce given attainments.

team confidence Shared beliefs of team members about the ability of the team to perform successfully.

Self-confidence in athletes is an athlete's belief or degree of certainty that he or she has the ability to perform successfully in sport.

1. DEFINITIONS OF SELF-CONFIDENCE

There are many terms used in the literature that relate to self-confidence. Self-efficacy is defined as beliefs in one's capabilities to organize and execute the courses of action required to produce given attainments. Self-confidence is the belief that one can successfully execute a given activity, and the sport-specific term sport-confidence has been used to describe an athlete's belief or degree of certainty that he or she has the ability to perform successfully in sport.

Research indicates that sport-confidence is multidimensional, or that confidence in athletes may be subdivided into specific types of confidence. Sport-confidence in physical skills and training (termed SC-physical skills and training) is defined as an athlete's belief or degree of certainty about his or her ability to execute the physical skills necessary to perform

©2004 Elsevier Inc.
All rights reserved.

successfully in sport. SC-cognitive efficiency is defined as an athlete's belief or degree of certainty that he or she can mentally focus, maintain concentration, and make effective decisions to perform successfully in sport. SC-resilience is an athlete's belief or degree of certainty that he or she can regain focus after performance errors, bounce back from performing poorly, and overcome doubts, problems, and setbacks to perform successfully in sport. Thus, self-confidence in athletes is the summation and integration of not only athletes' beliefs in their abilities to execute physical sport skills but also their beliefs in their abilities to be mentally efficient and resilient in relation to the unique competitive demands of competitive sport.

Although self-confidence is typically defined as a belief system (beliefs about one's ability to succeed as an athlete), it also can be viewed and assessed as an emotion. Research with athletes shows that "confident" is a positive emotion described as optimal to enhance sport performance. Similarly, "uncertain," which denotes a lack of confidence, is a negative emotional feeling described by athletes as very dysfunctional for their performance. Thus, self-confidence in athletes is a belief as well as an emotional feeling that significantly influences how they think and feel and thus perform.

Overconfidence is a popular term often confused as meaning that athletes have too much confidence. However, true self-confidence involves believing one can be successful, with the realistic understanding of the effort and perspective needed to be successful. From this perspective, athletes can never have too much true self-confidence. Thus, overconfidence is not really a problem with confidence, but rather it is a sense of complacency about the need to train and exert effort to be successful. Overconfidence is an overestimation of one's abilities and need for expended effort and an underestimation of the situational demands that one will face. Feelings of complacency have been identified by athletes as extremely dysfunctional to their performance.

Certain terms are sometimes confused, or assumed to be synonymous with, self-confidence. Perceived competence and perceived ability refer to athletes' perceptions about how much ability or competence they possess in a certain achievement domain (e.g., soccer). However, perceived ability/competence focuses on the skills that athletes have, whereas self-confidence focuses on athletes' beliefs about what they can do (e.g., perform successfully) with the skills that they have. Self-confidence, therefore, is an athlete's belief about how he or she is going to do in competition, rather than just a perception of how good he or she is as an athlete. Optimism has become a critical personality construct in relation to health and well-being but differs from self-confidence in that optimism is a tendency to expect the best possible outcome or dwell on the most hopeful aspects of a situation. Although different, optimism and self-confidence are related because confidence results from productive thinking habits that have become automatic and natural. Clearly, dispositional optimism is related to self-confidence due to the productive perceptions inherent in each. Other related terms represent similar "self" constructs, such as self-esteem and self-concept. Self-esteem pertains to an individual's personal perception of worthiness. Thus, athletes may have self-confidence but not self-esteem (and vice versa). Self-concept is more of an overall comprehensive view of oneself and incorporates both self-confidence and self-esteem.

In summary, although many constructs are similar and related to the notion of self-confidence, self-confidence is uniquely defined as a belief or self-perception about one's ability to perform successfully. For athletes, this involves believing one has the critical components for sport success, including physical skills and training, cognitive efficiency to focus and make effective decisions under competitive stress, and the ability to be resilient in responding to obstacles and poor performances.

2. IMPORTANCE OF SELF-CONFIDENCE TO ATHLETES

The influence of self-confidence on athletes is one of the most intriguing topics in sport psychology. Athletes often attribute "choking" and performance slumps to "losing" their confidence. Research has strongly supported that higher levels of self-confidence are related to success in sport. The specific ways in which self-confidence helps athletes to achieve success are presented in the following sections.

2.1. Influence of Self-Confidence on Thoughts and Emotions

Confident athletes are more cognitively efficient and emotionally adaptive than less confident athletes. Confident athletes have been shown to engage in more productive (personally controllable) attributions for their successes and failures, which serves to enhance their motivation. That is, when confident athletes fail,

they attribute their failure to unstable and personally controllable factors, which means their motivation remains stable and strong because they believe they can succeed in the future. Conversely, when athletes low in confidence fail, they attribute their failure to stable factors that are largely unchangeable (e.g., "I'm not any good"). Self-confidence also facilitates concentration or a task focus; research on basketball players showed that more confident players were more likely to narrow attention effectively and less likely to experience attentional overload and reduced attention. Athletes low in confidence tend to focus more on themselves and their perceived inadequacies and thus are unable to narrow attention to focus efficiently on performance demands. Higher self-confidence in athletes is also associated with more positive perceptions of arousal and anxiety because confidence provides a productive belief system in which negative emotions are reframed and viewed as necessary and facilitative. Athletes who lack confidence view their anxiety as negative and debilitating to their performance, and they lack the personal coping resources to deal productively with anxiety.

Confident athletes use more problem-focused coping strategies, as opposed to emotion-focused coping strategies. This again indicates an efficient and skillful use of cognitive resources in that coping should be problem focused to address important task issues as opposed to heightened emotionality, which does not directly address the problem and a solution. Confidence athletes also engage in more mastery imagery and have better imagery ability than less confident athletes. Imagery has been shown to facilitate elite status in sport and is an important mental skill for athletes. Finally, self-confidence elicits adaptive emotional states for athletes, whereas a lack of confidence relates to anxiety, depression, and dissatisfaction.

2.2. Influence of Self-Confidence on Behavior and Performance

Research has supported that self-confidence is positively related to performance in athletes. This has been shown in tennis, marathon running, baseball, basketball, field hockey, soccer, wrestling, track and field, gymnastics, and swimming. Research has also identified the behavioral mechanisms by which confidence enhances performance. Self-confidence increases athletes' effort and persistence in the face of obstacles, so these critical achievement behaviors along with enhanced cognitive and emotional efficiency serve to enhance the performance of athletes. Interestingly, laboratory studies have shown that manipulating levels of confidence in individuals on motor performance tasks (arm wrestling, weight lifting, and muscular endurance tasks) actually affects the amount of effort and persistence engaged in by the participants. If participants in these studies were convinced that they were stronger or more fit than their opponents, they would increase their effort and persistence. Conversely, if they were convinced that they were weaker or less fit than their opponents, they demonstrated less effort and persistence. Thus, athletes who lack confidence tend to exert less effort and persistence in the face of obstacles and failures, which is clearly detrimental to their performance and development.

As one might expect, the effects of performance on self-confidence are even stronger than the effects of confidence on performance. That is, self-confidence helps athletes perform better, but successful performances are even stronger predictors of athletes' subsequent levels of self-confidence. This is discussed further in the next section.

3. SOURCES OF SELF-CONFIDENCE IN ATHLETES

Most research in sport psychology on the sources of self-confidence has been conducted within the parameters of Bandura's self-efficacy theory. Bandura asserts that there are four sources of self-efficacy: (i) enactive mastery experiences, such as success; (ii) vicarious experiences, such as social comparison with others or watching successful models; (iii) verbal persuasion or the social influence of evaluative feedback, expectations of others, self-talk, and positive imagery; and (iv) physiological and affective states, such as being psyched up for competition or fatigued and listless, both of which influence subsequent levels of confidence in athletes. All four sources have been shown to be predictive of confidence in sport and motor activities, with enactive mastery experiences (or success) emerging as the strongest source of confidence, as predicted by theory.

A research line in sport-confidence has identified nine sources of confidence important to athletes in competitive sport (Table I). The nine sources of self-confidence for athletes are divided into three broad domains: achievement, self-regulation, and social climate. Achievement sources of confidence include mastery and demonstration of ability. These sources are similar to Bandura's source of enactive mastery

TABLE I
Sources of Sport Confidence in Athletes

Source	*Confidence derived from . . .*
Achievement	
Mastery	Mastering or improving your personal skills
Demonstration of ability	Showing off your skills to others or demonstrating more ability than your opponent
Self-regulation	
Physical/mental preparation	Feeling physically and mentally prepared with an optimal focus for competition and performance
Physical self-presentation	Perceiving that you physically look good
Social climate	
Social support	Perceiving support and encouragement from significant others in sport, such as coaches, teammates, and family
Vicarious experience	Watching others, such as teammates, perform successfully
Coach's leadership	Believing your coach is skilled in decision making and leadership
Environmental comfort	Feeling comfortable in a competitive environment
Situational favorableness	Feeling that the breaks or the nature of a situation are in your favor

experiences. However, for athletes in sport, important enactive mastery experiences that enhance their self-confidence involve mastering or improving their skills as well as demonstrating greater ability than their opponent (which often translates into winning). Thus, although winning is an important achievement source of confidence for athletes, they may also develop confidence from other types of success (e.g., mastery or skill improvement).

The second domain of sources of self-confidence for athletes is self-regulation, which includes physical/mental preparation and physical self-presentation. Feeling physically and mentally prepared with an optimal focus for competitive performance is a critical source of confidence for athletes. Research and a great deal of experiential evidence from successful athletes indicate that precompetitive routines, rituals, and focus plans are critical predictors of how well athletes perform, especially in championship or pressure situations. Physical conditioning has been rated by coaches as the most important way to develop self-confidence in athletes.

The third domain of sources of self-confidence for athletes is social climate, with five sources representing aspects of the physical and social environment that serve to enhance the confidence of athletes. Clearly, perceiving support and encouragement from others (social support) is an important source, as is vicarious experience, in which athletes (particularly young athletes) gain confidence to try skills after watching a teammate or friend perform them successfully. Very unique to the competitive sport environment is the source of coach's leadership, whereby athletes derive feelings of confidence from perceiving that their coaches are skilled leaders and effective decision makers. The final two sources in the social climate domain are also very unique to the competitive sport environment. Environmental comfort is often identified by athletes as a source of confidence gained from feeling comfortable in a particular competitive venue. The "home advantage," or research finding that home teams win more than 50% of their competitions played under a balanced home-and-away schedule, is often cited as a source of confidence for athletes. Situational favorableness involves gaining confidence by feeling that factors within the situation are in one's favor. For example, the popular notion of psychological "momentum" refers to athletes' perceptions that something has occurred that increases their probability of success, which typically creates a surge of confidence.

For individual sport collegiate athletes, the top five sources of self-confidence were physical/mental preparation, social support, mastery, demonstration of ability, and physical self-presentation. Physical self-presentation and social support have been shown to be more important sources of confidence for female athletes, indicating that body image along with social approval from others are learned to be important for females when competing in sport. Female athletes learn that maintaining socially defined feminine qualities affords them social acceptance and approval due to the emphasis placed on how women look in society. For team sport athletes, the top five sources of self-confidence were mastery, social support, physical/mental preparation, coach's leadership, and demonstration of ability. Demonstration of ability was a more important source of confidence for male athletes, indicating that male athletes typically experience extreme normative social comparison expectations with regard to sport prowess and participation.

Highly confident athletes tend to use personally controllable sources to nurture their confidence, including physical/mental preparation and mastery. Athletes with less confidence tend to base their confidence on less controllable sources, including environmental comfort and social support. Research on children and young athletes shows that frequency and intensity of negative feedback from adults and an emphasis on unrealistic goals and expectations can lower confidence levels. This is because young children base their confidence more on feedback from adults than do adolescents and adults. Overall, the sources of confidence that athletes develop and use are important determinants of their ultimate levels of self-confidence, and particularly the stability of their confidence over time. Research supports that athletes gain self-confidence when they achieve mastery goals; engage in effective self-regulation of cognitions, emotions, and behavior; and train and compete in a competitive environment that is supportive, challenging, comfortable, and motivating.

4. SELF-CONFIDENCE WITHIN SPORT TEAMS

People involved in sport understand that although self-confidence is an individual perception, there is also an important sense of team, or collective, confidence that is defined as the shared beliefs of team members in the ability of the team to perform successfully. Clearly, individual team members may be very confident about their individual abilities to perform successfully, but at the same time they may have little or no confidence in the team's ability to succeed. Team confidence is particularly important for highly interactive teams (e.g., volleyball, hockey, and basketball teams) as opposed to coacting teams (e.g., track and field or golf teams).

Research found that team confidence was a better predictor of team outcomes for ice hockey than the aggregate of individual levels of confidence of team members. Also, team confidence increased after winning and decreased after losing, whereas individual self-confidence was not influenced by team outcomes. Teams with high levels of confidence have also been shown to be more cohesive in terms of team members' feelings of involvement with the team's productivity. Finally, athletes' perceptions that there is strong leadership within the team was related to overall team confidence, particularly for teams that were unsuccessful the previous season. This finding indicates that the confidence of teams coming off losing seasons may be positively enhanced if leadership abilities within the team are apparent to athletes on the team.

5. MEASURING SELF-CONFIDENCE IN ATHLETES

Self-confidence in athletes has been measured in many ways, depending on the conceptual perspective of the researcher. Several inventories have been developed to measure various aspects of confidence in athletes. The Trait Sport-Confidence Inventory and State Sport-Confidence Inventory have been used to measure dispositional sport-confidence and "right now" (often precompetitive) sport-confidence in athletes, respectively. The Sport-Confidence Inventory has been developed to assess not only total sport-confidence in athletes but also the various types of sport-confidence shown to be relevant to athletes in competitive situations (SC-physical skills and training, SC-cognitive efficiency, and SC-resilience). In addition, the Sources of Sport-Confidence Inventory is used to measure the nine sources of sport-confidence discussed previously.

Using a self-efficacy conceptual approach, researchers measure self-confidence in athletes specific to a particular task (e.g., volleyball serving as opposed to confidence in overall volleyball performance abilities). Self-efficacy measures can be hierarchical or nonhierarchical. Hierarchical self-efficacy measures are typically constructed by listing a series of tasks that vary in difficulty, complexity, or stressfulness. Athletes are asked to designate (yes or no) the tasks they believe they can perform (efficacy level). For each task designated as yes, athletes rate their degree of certainty (efficacy strength) that they can execute it, from total uncertainty to total certainty (scales range from 0 to 10 in 1-unit increments or 0 to 100 in 10-unit increments). Nonhierarchical self-efficacy measures provide a list of subskills needed to perform successfully in a specific sport (e.g., serving, passing, hitting, blocking, back row defense, and setting in volleyball). Athletes are asked how sure they are that they can successfully perform each specific task at a certain level (e.g., 1, 2, 3, . . . , 10 out of 10 times). This assesses the level of self-efficacy for each task. Athletes are then asked how certain they are that they can perform each task at each level (usually 0–100%). This assesses strength of self-efficacy, or how sure athletes are that they can attain each level of each task.

6. STRATEGIES TO ENHANCE SELF-CONFIDENCE IN ATHLETES

Self-confidence may be enhanced in athletes by strengthening the sources of confidence that are important to athletes and success in sport.

6.1. Quality of Physical Training and Perceived Achievement

The most obvious point about developing confidence in athletes is that self-confidence must be earned through the development of physical skill and fitness to serve as a basic prerequisite for performance success. Quality training to hone skills and fitness development are foundational achievements that serve as the basis for confidence building. Athletes have rated physical preparation as one of their top sources of confidence, and successful world-class athletes have particularly emphasized the importance of quality training for their confidence and performance success. Coaches and sport psychology consultants should identify training methods that enhance confidence, such as simulating pressure situations and creating unexpected obstacles to train adaptability in athletes.

Another basic foundation for athletes' confidence is their perception of their past achievement. The key here is the emphasis on perceptions, with the point being that athletes must engage in productive attributional assessments of their achievement attempts and perceptions of success based on controllable factors (e.g., appropriate goal setting and evaluation). Higher levels of confidence in athletes have been shown to be related to focusing on sources or strategies that athletes personally control, such as mastery of physical skills, whereas lower levels of confidence were related to uncontrollable socially comparative or environmental sources. Coaches and sport psychology consultants should help athletes define success in personally controllable ways, such as through the use of performance and process goals.

6.2. Self-Regulation

Self-regulation has been identified as a significant source of confidence for athletes, and it is the central focus of interventions in sport psychology. To self-regulate in sport means that athletes must effectively manage their behaviors, thoughts, and feelings in relation to competitive challenges and demands. Goal mapping, imagery, and self-talk are the three basic self-regulatory tools advocated by sport psychologists to enhance confidence.

The focused and persistent pursuit of goals serves as a basic regulator of human behavior; thus, it seems obvious that goal mapping is an important strategy to enhance the confidence and performance of athletes. A goal map is a personalized plan for an athlete that contains various types of goals and goal strategies as well as a systematic evaluation procedure to assess progress toward goals. Although athletes typically talk about the motivating qualities of outcome goals (winning or making the Olympic team), research indicates that performance and process goals are better sources of confidence. These types of goals seem to enhance confidence by allowing athletes to feel a sense of control over their performance and also by providing a specific task focus for productive concentration.

Imagery, or visualization, is assumed to enhance confidence in two ways. First, visualizing one's self successfully performing in sport may, in an individualized vicarious sense, provide a perfect mental model for performance. Imagery in this sense serves a cognitive function to improve or refine the physical reproduction of a motor skill. Second, imagery may enhance confidence through its motivational effect on behavior, creating a productive task focus of "what to do," and/or by creating positive emotion to energize confidence and performance. Researchers have shown that highly confident athletes tend to image the mastery and emotion associated with sport competition more so than less confident athletes. Thus, imagery interventions to enhance confidence should focus not only on the cognitive reproduction of skill execution but also on the emotional, attentional, and motivational aspects of performing in sport. Research has demonstrated that imagery training enhances sport-confidence in athletes.

Managing thoughts, or self-talk, is advocated by applied sport psychology sources as an important key for self-confidence. Athletes who effectively manage their self-talk are, in essence, controlling how they think as opposed to letting external events dictate how they think. This is essential due to the social cognitive nature of sport-confidence. That is, confidence is a belief system—a set of beliefs about one's ability to be successful in sport. Self-talk is representative of this belief system because what athletes say to themselves is what they believe. The difficulty in managing self-talk lies in the automaticity and invisibility of athletes' thoughts that make up their belief systems.

These belief systems are a product of past social interactions that have reinforced certain ideas about their abilities. These belief systems directly influence, and in fact define, one's self-confidence.

There are several variations of self-talk interventions that may serve to enhance confidence, including negative thought-stopping, countering, reframing, and the use of affirmations. Negative thought-stopping teaches athletes to recognize negative thoughts, self-monitor across time to learn individualized thought patterns in response to social stimuli, and habituate personal cues that allow them to stop negative thoughts from becoming internalized into their belief systems. Countering involves self-debate to understand how irrational and inappropriate beliefs create negative thinking, and this intervention has been shown to enhance confidence in athletes. Reframing involves choosing to view events in alternative ways, typically in a more productive way, and affirmations are a form of productive self-statements repeated over and over to internalize a belief or attitude. All these interventions focus on the development and internalization of a positive belief structure about one's self; thus, all serve to enhance confidence.

Overall, self-regulation is based on the idea of choice. Athletes have the volition, or will, to choose how they act, think, and feel. Social conditioning clouds this perception and seduces individuals into believing that their thoughts and feelings are controlled by their social and physical environments. The key to consistent, healthy confidence in sport is the exercise of personal volition through self-regulation. Although this is difficult, particularly in the organizational culture of sport, it is possible to learn and enhance self-regulatory skills to maintain optimal self-confidence in sport.

6.3. Social Climate

Self-regulation has been hailed as an important and successful strategy to enhance confidence, but coaches and sport psychology consultants must keep in mind that human behavior always takes place within a social climate. Intervention plans designed to enhance the motivational climate of educational settings have increased the motivation and confidence of learners. Social climate factors that seem to influence confidence are leadership style, types of goal recognition and evaluation, social feedback sources and types, and availability and characteristics of models.

Confidence is enhanced when athletes are part of a productive and congruent social support network that can be facilitated through team-building activities. The social influence of coaches' leadership styles and feedback patterns have been shown to be important influences on athletes' perceptions of competence and confidence. Based on the importance of vicarious experience as a source of confidence, athletes should be provided effective models for not only skill execution but also achievement behaviors such as effort and persistence. Although environmental comfort emerged as a significant source of confidence for athletes, teaching athletes to rely on such external and uncontrollable factors to enhance their confidence seems problematic. Athletes should learn to base their confidence more on self-regulatory control of perceptions, emotions, and behavior so that they are not subject to the inconsistent type of self-confidence that is dependent on social and environmental confidence builders.

See Also the Following Articles

Arousal in Sport ■ Attention and Concentration Training in Sport ■ Intrinsic and Extrinsic Motivation in Sport ■ Psychological Skills Training in Sport ■ Successful Athletic Careers

Further Reading

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.

Feltz, D. L., & Lirgg, C. D. (2001). Self-efficacy beliefs of athletes, teams, and coaches. In R. N. Singer, H. A. Hausenblas, & C. M. Janelle (Eds.), *Handbook of sport psychology* (2nd ed., pp. 550–565). New York: Wiley.

George, T. R. (1994). Self-confidence and baseball performance: A causal examination of self-efficacy theory. *Journal of Sport & Exercise Psychology*, *16*, 381–399.

Grove, J. R., & Heard, N. P. (1997). Optimism and sport confidence as correlates of slump-related coping among athletes. *The Sport Psychologist*, *11*, 400–410.

Jones, G., & Swain, A. B. J. (1995). Predisposition to experience debilitative and facilitative anxiety in elite and nonelite performers. *The Sport Psychologist*, *9*, 201–211.

Moritz, S. E., Hall, C. R., Martin, K. A., & Vadocz, E. (1996). What are confident athletes imaging? An examination of image content. *The Sport Psychologist*, *10*, 171–179.

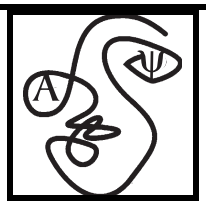
Vealey, R. S. (2001). Understanding and enhancing self-confidence in athletes. In R. N. Singer, H. A. Hausenblas, & C. M. Janelle (Eds.), *Handbook of sport psychology* (2nd ed., pp. 550–565). New York: Wiley.

Vealey, R. S., Hayashi, S. W., Garner-Holman, M., & Giacobbi, P. (1998). Sources of sport-confidence:

Conceptualization and instrument development. *Journal of Sport & Exercise Psychology*, *20*, 54–80.

Vealey, R. S., & Knight, B. J. (In press). Conceptualization and measurement of multidimensional sport-confidence. *Journal of Sport & Exercise Psychology*.

Zinsser, N., Bunker, L., & Williams, J. M. (2001). Cognitive techniques for building confidence and enhancing performance. In J. M. Williams (Ed.), *Applied sport psychology: Personal growth to peak performance* (pp. 284–311). Mountain View, CA: Mayfield.



Self-Control

Kathleen D. Vohs
University of British Columbia, Vancouver, British Columbia, Canada

Roy F. Baumeister
Florida State University, Tallahassee, Florida, USA

1. Definitions and Conceptual Considerations
2. Research Paradigms and Findings
3. Self-Control Process Models
4. Domains of Self-Control
5. Behavioral Change
6. Individual Differences
7. Summary
   Further Reading

GLOSSARY

delay of gratification The idea that self-control is important in terms of being able to delay indulging in impulses or succumbing to temptation. Longer delays have greater benefits.

ego depletion A state of depletion of regulatory resources, typically associated with subsequent self-control failure.

emotional control Attempts to change the valance, intensity, or duration of one's emotional reactions. Most emotion control attempts are aimed at ameliorating negative affect.

Low Self-Control Scale A trait scale created to measure the effects of low self-control abilities on criminality.

mental control Attempts to change the content or frequency of one's thoughts.

regulatory resource model The idea that self-control is governed by a finite supply of energy that is shared among all varieties of self-control.

self-control The intentional, conscious process of changing one's behaviors or responses.

Self-Control Scale by Tangney and Baumeister A trait scale that measures generalized self-control abilities in the domains of self-discipline, reliability, work ethic, healthy habits, and deliberate action.

self-control subscale of the California Psychological Inventory A trait scale created to measure ebullient and joyful relinquishments of self-control.

self-discipline People's actions to achieve deliberate goals by way of self-improvement.

self-regulation Modification, alteration, or guidance of one's responses (be they automatic or conscious) to achieve a desired goal or end state.

TOTE model shorthand term for the test–operate–test–exit cybernetic loop. Self-control is conceptualized as a cyclical set of processes in which one moves increasingly closer to the goal ("tests" and "operates") and then finally "exits" the loop when the goal is reached.

The ability to regulate the self has been deemed "the hallmark of human functioning" because it constitutes a vital human skill. Mental and physical health rely heavily on the capacity to effectively and consistently exert control over incipient responses, often by inhibiting an undesirable response and then substituting a desirable response. Failures to self-regulate underlie health problems ranging from risky sexual practices to poor nutrition and improper diet, child abuse, overspending, and all manners of addictive problems. Accordingly, understanding how, why, and under which conditions people fail at self-control is a goal that has taken on paramount importance in recent decades.

©2004 Elsevier Inc.
All rights reserved.

1. DEFINITIONS AND CONCEPTUAL CONSIDERATIONS

The term self-control is used to indicate the intentional, conscious process of suppressing or inhibiting one's behaviors or responses. Self-discipline is a term that is subsumed under the concept of self-control. It is used to indicate people's actions to achieve deliberate goals that are often in the service of self-improvement. It is useful to distinguish the term self-control from the term self-regulation. Self-regulation is a broad term used to indicate the modification, alteration, or guidance of one's responses (be they automatic or conscious) to achieve a goal. The focus of this article is self-control, which means that all the findings discussed here could also be considered self-regulation research but only some of which are self-discipline research.

2. RESEARCH PARADIGMS AND FINDINGS

This section outlines the major processes and domains of self-control research. The aim of this section is to demonstrate the basic processes that are thought to underlie self-control attempts as well as to illustrate the breadth of responses that are governed by self-control.

3. SELF-CONTROL PROCESS MODELS

3.1. Cybernetic Model

Charles Carver and Michael Scheier developed the cybernetic model for self-control when they introduced the idea of self-control as occurring within a test–operate–test–exit (TOTE) loop. People enter the TOTE loop when they establish a goal. The first action, test, refers to the comparison between the current and goal state. Assuming there is a mismatch between current and desired state, the person engages the operate function of the model in which responses are enacted to close the gap. The person tests again and, depending on whether he or she deems the goal to be met or whether more work is needed, the person either returns to the operate phase or exits the loop.

One important idea from Carver and Scheier's TOTE model is that emotions reflect progress toward a goal: Positive affect is often a cue that a person is moving toward the goal and negative affect signals movement away from the goal.

3.2. Regulatory Resource Model

The regulatory resource model (or ego depletion model) posits that the ability to control oneself is governed by a finite resource that is shared among all types of self-control faculties. (Readers may notice that the resource model specifies the operate function in a TOTE model.) Any one act of self-control leaves the person less successful for a subsequent act (within a circumscribed time period) because fewer resources are available to put toward the second self-control action. This is known as a state of ego depletion, which reflects an inability to muster enough regulatory resources to meet a specific self-control demand. Roy Baumeister, Kathleen Vohs, and colleagues posited that the responses involved in emotion regulation, mental control, impulse intervention, and behavioral guidance all use this resource and, consequently, levels of the resource diminish with each act of self-control.

Two examples of studies from the resource model are relevant to applied psychology because they involve dieting and impulsive spending. In these studies, people were instructed to control some aspect of their behavior (e.g., attentional or emotion responses). This was intended to deplete regulatory resources for these participants, relative to participants who were not instructed to engage in self-control. Subsequent behaviors in the realms of eating ice cream and unplanned purchasing patterns were assessed. People who had earlier controlled themselves in one domain were less able to control themselves with respect to eating or impulsive spending.

3.3. Delay of Gratification Paradigm

The third model of self-control centers on the ability to delay gratification. For more than 40 years, research by Walter Mischel and colleagues has illuminated the importance of curbing frustration, stifling unwanted responses, and overcoming temptation for goal achievement, psychological growth, and well-being. Typical studies testing the delay of gratification perspective focus on children approximately 3 or 4 years old, who are seated in front of a table on which sits an inviting treat. The child is told that if he or she does not eat the treat (e.g., one marshmallow) in front of him or her now then he or she will be given a larger

treat (e.g., two marshmallows) later. Delay of gratification is measured in terms of whether and how long the child can wait and not eat the available treat.

The adaptiveness of being able to delay gratification has been well demonstrated. Children who at 4 and 5 years of age had better delay of gratification skills had higher SAT scores, got along better with others, and dealt with life stressors more effectively. These data speak to the long-lasting benefits of early self-control and the importance of self-restraint skills.

4. DOMAINS OF SELF-CONTROL

4.1. Mental Control

It is no small feat to rid oneself of an unwanted thought, and research on thought suppression and mental control has illuminated the sheer difficulty of such intents. Psychologist Daniel Wegner built an important line of research around asking people to suppress thoughts of a white bear. People given such instructions find it incredibly difficult to accomplish this goal; moreover, when the suppression task ended, people experienced a surge in white bear thoughts. This effect has been replicated in studies on suppressing thoughts about sex, secret romantic relationships, and depressive thoughts.

4.2. Emotion Control

There are approximately six different categories of affect regulation: getting into, out of, or maintaining a positive or negative emotion. Most affect control efforts, however, are aimed at just one of these categories: ameliorating negative emotions. Researcher James Gross has conceptualized emotion regulation as occurring both pre- and post-emotional stimulation. Situation selection involves altering emotional state by choosing specific types of people, places, and objects to change one's mood. Situation modification strategies are efforts to alter a situation to change its emotional effects. Attentional deployment isolates aspects of the situation to manage emotional state. For instance, people who are scared of flying may chose to watch the in-flight movie. Cognitive change involves reconstruing the situation in alternative ways to lessen the impact of the emotional stimulus. Last, response modulation refers to the direct control of emotional responses after they are initiated.

5. BEHAVIORAL CHANGE

Recent models of behavioral self-control have highlighted the importance of distinguishing the start of behavioral change from the maintenance of behavioral change. Psychologist Alexander Rothman found that what people want to get out of a behavioral change program determines whether they make a behavioral change, but satisfaction with progress determines whether the change will be maintained. A person considering dieting will therefore consider how he or she will look in a new bathing suit when the diet is complete. If this expectancy is positive, the person is likely to start a diet. Once started, however, whether the person is satisfied with the perceived progress predicts continuation of the program. This analysis of behavioral change has been a powerful tool to researchers and clinicians for determining how to motivate people to begin self-change programs and to stay on track.

5.1. Adherence to Moral Rules

Adhering to cultural rules and living up to a society's values is an important life ability that also requires self-control. Research suggests that children who are toddler aged have little self-control, and that self-control begins to emerge at approximately 5–10 years of age in response to punishment fears. However, Grazyna Kochanska and colleagues have shown that in children of toddler and early school age, inhibitory control predicts children's conscience, including moral behavior, moral cognitions, and moral identity and self-concept. Judith Rich Harris suggested that by preadolescence and adolescence, children demonstrate self-control primarily in response to interpersonal and peer isolation cues. By the time people are in their early 20s, they are behaving in a self-controlled manner that is primarily in response to internalized cultural values.

6. INDIVIDUAL DIFFERENCES

This section reviews trait scales of self-control. They have been derived from different histories and backgrounds and thus have provided the field with a variety of options for measuring different types and manifestations of self-control.

The oldest of the three scales reviewed is the self-control (Sc) subscale from the California Psychological Inventory (CPI). The 38-item scale was initially created

to measure "the kind of joyful, ebullient abandonment of restraint that one sees at certain times such as attendance at a carnival." Accordingly, Harrison Gough and colleagues set about to measure "impetuosity, high spirits, caprice, and a taste for deviltry" and hence the Sc scale of the CPI was born. This scale gained followers in the 1960s and 1970s, but it did not enjoy a very popular run due to some psychometric problems. Many of the items on the CPI Sc scale are too varied (overly heterogeneous) and lack validity. Moreover, the scope of the scale is too narrow.

The socialization subscale of the CPI is also widely used in applied settings. It has 46 items and is best known as a scale to measure impulse restraints, revealing whether a person is reliable, thorough, industrious, well socialized, and rules and directions are followed without problem.

The Five Factor Personality Model, popularized by Paul Costa and Robert McCrae, has also been used to predict self-controlled behaviors. The personality dimension of conscientiousness best approximates a self-control scale, although extraversion has also shown some predictive abilities. Conscientious individuals are better able to make and carry out long-term plans, which means that they have better self-control, especially regarding the practical aspects of self-control (e.g., planning, being reliable, and keeping organized).

From the study of criminology, Harold Grasmick and colleagues set out to operationalize ideas put forth in Michael Gottfredson and Travis Hirschi's book, *A General Theory of Crime*, which conceptualized aggression and criminality as poor self-control. The resulting Low Self-Control Scale has been used by criminologists to study crime, aggression, and delinquency, and it has proven its mettle by showing that people who score low in self-control are more likely to have problems abiding the law and following codes of conduct.

The Low Self-Control Scale contains 24 items, including statements such as

- I seem to have more energy and a greater need for physical activity than most other people my age.
- I will try to get things I want even when I know it's causing problems for other people.
- When things get complicated, I tend to quit or withdraw.
- I like to get out and do things more than I like to read or contemplate ideas.
- I'm not very sympathetic to other people when they are having problems.
- I often act on the spur of the moment without stopping to think.
- I frequently try to avoid projects that I know will be difficult.
- I'm more concerned with what happens to me in the short run than in the long run.

Participants are asked to indicate their agreement with these statements as descriptions of "the type of person you are." The scale is a unidimensional construct.

The most recently developed scale is the Self-Control Scale by June Tangney and Roy Baumeister. Their scale was created to measure varieties of self-control abilities and failures that have psychological and practical implications, such as breaking habits, resisting temptation, and maintaining self-discipline. This scale has a 36-item long form and a 10-item short form. The long form has five subscales: self-discipline, deliberate/nonimpulsive action, healthy habits, work ethic, and reliability.

The Self-Control Scale contains 36 items, to which participants are asked to assign a numeric value (where 1 = "not at all like me" and 5 = "very much like me") that "best represents what you believe to be true about yourself for each question." Representative items include

- I have a hard time breaking bad habits.
- I am lazy.
- I say inappropriate things.
- I never allow myself to lose control.
- I do certain things that are bad for me, if they are fun.
- People can count on me to keep on schedule.

People who score high on this scale get better grades in college, report less psychopathology, have higher self-esteem, are less likely to engage in binge eating or alcohol abuse, have a more secure attachment style, report better relationships with family and friends, and show fewer negative emotional reactions.

7. SUMMARY

This article discussed the core elements of self-control, including theoretical, empirical, and practical treatments of it. Self-control is one of the greatest human abilities because it gives us the flexibility to (psychologically) move beyond the immediate physical setting and orient our responses toward abstract, higher order goals. It is likely that self-control abilities are one of the key dimensions that allowed humans to advance evolutionarily, given its advantages for securing safety and increasing reproductive possibilities.

The psychological literature is clear on the importance of self-control for overcoming immediate urges, controlling thoughts, modifying emotional states, and altering behavior. Having good self-control, at a personality level, is related to better outcomes in the realms of physical health, mental well-being, interpersonal relationships, and cultural rule-abiding (i.e., criminality). In short, there are very few realms in which having good self-control is not a benefit to the self.

See Also the Following Articles

Emotion

Further Reading

Baumeister, R. F., Heatherton, T. F., & Tice, D. M. (1994). *Losing control: How and why people fail at self-regulation*. San Diego: Academic Press.

Baumeister, R. F., & Vohs, K. D. (In press). *Handbook of self-regulation research*. New York: Guilford.

Carver, C. S., & Scheier, M. F. (1981). *Attention and self-regulation: A control theory approach to human behavior*. New York: Springer-Verlag.

Faber, R., & Vohs, K. D. (In press). To buy or not to buy? Self-control and self-regulatory failure in purchase behavior. In R. F. Baumeister, & K. D. Vohs (Eds.), *Handbook of self-regulation research*. New York: Guilford.

Gross, J. J. (1999). Emotion and emotion regulation. In L. A. Pervin, & O. P. John (Eds.), *Handbook of personality: Theory and research* (2nd ed., pp. 525–552). New York: Guilford.

John, O. P., & Srivastava, S. (1999). The big five trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin, & O. P. John (Eds.), *Handbook of personality* (2nd ed., pp. 102–138). New York: Guilford.

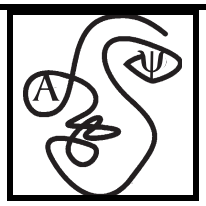
Mischel, W., Shoda, Y., & Peake, P. K. (1988). The nature of adolescent competencies predicted by preschool delay of gratification. *Journal of Personality and Social Psychology*, *54*, 687–696.

Sayette, M. A. (In press). Self-regulatory failure and addiction. In R. F. Baumeister, & K. D. Vohs (Eds.), *Handbook of self-regulation research*. New York: Guilford.

Tangney, J. P., Baumeister, R. F., & Boone, A. L. (In press). High self-control predicts good adjustment, less pathology, better grades, and interpersonal success. *Journal of Personality*.

Vohs, K. D., & Heatherton, T. F. (2000). Self-regulatory failure: A resource-depletion approach. *Psychological Science*, *11*, 249–254.

Wegner, D. M. (1994). Ironic processes of mental control. *Psychological Review*, *101*, 34–52.



Sexual Behavior

John D. Baldwin and Janice I. Baldwin
University of California, Santa Barbara, California, USA

1. Before Puberty
2. Puberty and Later
3. Early Sexual Problems
Further Reading

GLOSSARY

double standards The notion that society often treats men and women differently, expecting and accepting unequal behavior from the two genders.

observational learning The ability of people to learn about behavior by observing other individuals who perform the behavior; men tend to imitate men more than they do women, and women tend to imitate women more than they do men.

operant learning The ability of individuals to learn from the consequences of their own actions; behavior that is followed by rewards tends to become more frequent in the future, and behavior that is followed by punishment tends to become less frequent in the future.

Pavlovian conditioning The notion that stimuli that precede and predict the onset of a stimulus–response reflex become conditioned stimuli that have the capacity to elicit similar reflexive responses.

punishers Behavioral consequences that are "bad outcomes"; they suppress the behavior they follow, producing inhibitions about repeating the behavior.

reinforcers Behavioral consequences that are "good outcomes"; they strengthen (or reinforce) the behavior they follow.

sexual turn-ons Conditioned stimuli that have become associated with the sex reflex through Pavlovian conditioning; they are also called erotic stimuli and are capable of eliciting sexual arousal.

One of the best ways in which to understand human sexuality is to take a developmental approach, tracing sexual behavior from infancy onward. Babies are capable of basic sexual responses from the early days of life, and it is informative to study how more complex sexual behavior develops from the early responses of infants. This knowledge is of great practical value for all students of sexuality, including sex therapists and any individuals who want to improve their sexual experiences in life. The developmental approach allows one to begin with the most basic elements of sexual behavior and then add layers of extra information as one traces the emergence of sexual behavior from infancy to adulthood. Developmental theories are also well suited for showing how nature and nurture interact. The development of human sexuality involves complex mixes of biological and social inputs—both nature and nurture—mediated by multiple neural systems that operate at both conscious and unconscious levels, intertwining biologically based reflexes, learned responses, and various degrees of self-control. Naturally, it takes many years of sexual learning for a person to acquire all of the layers of complexity that are typical of adult sexuality. Not surprisingly, some people acquire far more layers of intricate sexual knowledge and behavior than do others.

1. BEFORE PUBERTY

Long before the onset of puberty, many infants and children have sexual experiences, including spontaneous

©2004 Elsevier Inc.
All rights reserved.

sexual arousal, erections, vaginal lubrication, self-stimulation, and sometimes sex play. Even though childhood sexual activities are less complex than the sexual conduct seen after the onset of puberty, they show that the sexual organs are well developed and operative early in life. People are not born as sexual "blank slates," as some early scholars once assumed. Before birth, various physiological mechanisms have developed to produce the sexual responses seen in early life.

The first sexual responses of babies are purely reflexive rather than consciously planned or intended. There are reflex mechanisms located in the lower half of the spinal cord that function from birth to produce vaginal lubrication, penile erection, and other basic sexual responses. Soon after being born, many boys have penile erections, and some girls have vaginal lubrication and clitoral tumescence. Ultrasound imaging of pregnant women reveals that male fetuses sometimes have erections in utero, several months before birth. Given current technology, it is difficult to measure the sexual responses of female fetuses, but there are indications of genital swelling and vaginal lubrication.

An infant's sex reflexes can be elicited by mild stimulation of the genitals. Penile erection in baby boys and vaginal lubrication in baby girls may occur when bed sheets or clothing rub the infants' genitals, when babies explore touching their genitals with their hands, or when other people fondle the infants' genitals. Prolonged touching can lead to sexual excitement, and a small percentage of children learn to masturbate to orgasm during the first year of life.

In some cultures, adults gently touch their babies' genitals as a way in which to pacify the infants, indicating that genital stimulation is pleasant for infants. In the U.S. culture, adults who stimulate children's genitals might be charged with child sexual abuse. However, infants often learn to touch their own genitals, even if they have never been touched by other people.

1.1. Early Operant Learning

It is very natural for children to use their hands to explore nearly anything that is within their reach, including parts of their own bodies. In this process of exploration, many infants discover that touching their own genitals brings pleasure. Technically speaking, genital stimulation is a "primary reinforcer" because nerves from the genitals are biologically connected to the "pleasure centers" in the brain and genital stimulation produces the rewarding feelings associated with sex. When infants first touch their genitals, they are merely exploring the world they can reach with their hands, with no intentions of discovering sexual pleasure. They do not innately know how to masturbate or produce sexual pleasure. However, their random explorations can produce positive reinforcers if they happen to touch their genitals, and these reinforcers increase the likelihood that, on later occasions, exploring infants will repeat those types of self-stimulation that were pleasurable in the past.

As children have repeated experiences with self-exploration and operant learning, some master increasingly effective ways of producing sexually pleasurable sensations. The types of genital self-touching that bring the most pleasure are reinforced and become stronger habits, whereas the types of self-stimulation that do not produce much pleasure become less frequent due to lack of positive reinforcement. As a result, some infants learn relatively effective forms of self-stimulation that adults may label with the term "masturbation" even though babies cannot verbalize or subjectively understand such words. Nevertheless, some infants learn how to stimulate themselves to orgasm before their first birthday.

Human sexual behavior is based on a combination of nature and nurture. Biological processes (i.e., nature) provide the infant with sex reflexes as well as the neural connections needed to experience sexual arousal, pleasure, and orgasm—and to learn from those experiences. Both personal experiences and social influences (i.e., nurture) are important as well. The power of nurture can be readily seen in the ways in which parents influence their children's operant learning by their responses to the children's early sexual exploration. Some parents recognize how natural it is for children to touch themselves, and they allow their children to explore many forms of self-touching. At the other end of the continuum are parents who punish their children for touching their genitals, perhaps by using a spanking, a firm grip on the arm, or a stern tone of voice. Parents who use strong punishment can suppress their children's masturbation and produce long-lasting inhibitions. Of course, many parents fall somewhere between the two ends of this continuum, allowing some pleasurable sexual experiences while giving various mixtures of punishment and inhibitions.

Parents who permit their children to explore and discover genital stimulation without punishment allow the children to learn masturbatory skills through the natural reinforcers of sexual pleasure, although they may also teach their children to restrict masturbatory activities to private locations such as bedrooms. On the other hand, many parents use considerable punishment

when they see signs of early sexuality; hence, many children stop masturbating during early childhood. Strong punishment may cause them to feel inhibited about masturbating, not only during childhood but also for many years later. The more punishment a child receives for masturbation, the stronger the inhibitions are likely to be, especially if the child is told that the masturbation is dirty, evil, or sinful.

Common sexual behaviors reported by mothers of 2- to 5-year-old children include masturbation, touching the mother's breasts (or other women's breasts), and trying to peer at adults or children when they are undressing or nude. As the children become older, their sexual behaviors gradually diminish. Mothers of 6- to 9-year-olds reported that their children primarily masturbated at home and attempted to look at people who are undressing or nude. By 10 to 12 years of age, the only common sexual behavior that mothers reported was an interest in the opposite sex.

1.2. Early Pavlovian Conditioning

Children's emotional responses to all aspects of sexuality result from the pleasurable conditioning that occurs as a side effect of reinforcement and punishment. Children who are allowed to explore their sexuality and discover sexual pleasure without punishment experience positive Pavlovian conditioning that minimizes inhibitions about sexuality. If parents avoid associating guilt and punishment with sex and explain that sex is a natural healthy part of life (especially when done in privacy), children can grow up feeling good about their sexuality. In contrast, children who have been punished for engaging in sexual activities may learn to associate sex with fear and anxiety through Pavlovian conditioning. In the future, doing sexual activities—or even thinking about them—can elicit negative emotions and inhibitions.

Childhood masturbation and Pavlovian conditioning can lead to the early development of sexual fetishes, although this is not common. A classic example in the sex literature is that of a man who developed a sexual fetish for rubber that could be traced to early childhood, when his mother had him sleep on a rubber sheet. He discovered pleasurable experiences when he accidentally rubbed his penis against the rubber bed sheet. Because he repeated this activity on many subsequent occasions, he developed a strong and long-lasting attraction to rubber objects (due to Pavlovian conditioning), and he began to collect all sorts of rubber items. Even though his mother often took his rubber toys away from him, he would collect more, often using them during masturbation. By adulthood, his positive conditioning for rubber had become strong enough to be called a "rubber fetish," and he enjoyed and defended this fetish as a special part of his sex life.

1.3. So What Is "Normal" or "Typical" During Childhood?

In 2001, Kaeser and colleagues studied kindergarten, first-grade, and second-grade children to determine the common or "normal" sexual behaviors at that age. Most common were talking among peers and close-age siblings about genitals and reproduction, showing one's genitals (if the peer or close-age sibling would reciprocate), playing "doctor," masturbating, imitating seductive behaviors such as kissing and flirting, and occasionally using "dirty" words or telling dirty jokes. A variety of behaviors were judged as possibly problematic, including a preoccupation with either masturbation or sexually explicit conversations with peers. Activities that were judged to be definitely problematic included having sexually explicit conversations with people who were notably different in age and repeatedly touching the genitals of others. Behaviors judged to need therapeutic intervention were acts such as oral, vaginal, or anal penetration of dolls, animals, or children and simulated intercourse with peers after removing clothing.

1.4. Double Standards

Many children are exposed to double standards about sexual activities. First, parents are twice as likely to punish their daughters for masturbation than to punish their sons, leading boys to believe that their parents are more accepting of sexual self-stimulation than girls believe their parents to be. Also, parents are three times more likely to tell their daughters that masturbation is sinful or morally wrong than to tell their sons. Thus, it is not surprising that girls and women often feel more uncomfortable than do boys and men about exploring sexual forms of self-pleasuring. This does not mean that women cannot overcome negative feelings about sexual self-stimulation, even those that were learned during early childhood. Sexual inhibitions are not "fixated" or "imprinted" during childhood experiences, as Freud and some early psychologists assumed. Sexual inhibitions can be reduced, through Pavlovian counterconditioning, by nearly any kind of pleasurable experiences that a person has associated with sex. A person with many sexual inhibitions can overcome

them, especially if he or she has a sexual partner who is very positive about sex as a natural and wonderful aspect of human experience and makes sure that both partners have very pleasurable experiences with sex.

Another double standard about childhood sexuality can be seen in the ways in which parents give names to their children's genitals. In our society, many parents teach their sons to name their genitals as a "penis" or some euphemism such as a "pee-pee." In contrast, parents often give their daughters less useful names for the vagina such as a "pocketbook that you'll open after you get married." Few girls are told about their clitoris and its name. Parents might fear that if they tell their daughters that they have a clitoris, the girls might ask, "So what's it for?" Many parents would be reluctant to answer that its only function is to produce sexual pleasure.

1.5. Early Observational Learning

During the early years of life, children often learn bits and pieces of information about sexuality by observing adults, television, movies, magazines, books, and early sex education programs. Children are often more attentive to sex information and learn more from observation than adults realize. The belief that observant children are not learning much about sex results from the fact that most children rarely talk about or act out the sexual behaviors they have observed. Why? Many children are punished if they act on their sexual observations, and this inhibits such behavior. In addition, children do not derive as much pleasure (or sexual reinforcement) as do teens and adults when they do act out sexual behaviors that they learned by observation because the pleasures of sexual stimulation are less intense before puberty than after puberty. This, in turn, helps to explain why prepubertal children cannot understand why teens and adults find sex so fascinating. Of course, some prepubertal children learn very little about sex because their parents prevent them from seeing sexual materials.

Parents are often quiet when their children are observing sexual information, as if they are hoping that their children are oblivious to the sexual implications of media images and innuendos. Nevertheless, there are so many modern television shows and Web sites that show blatantly sexual images that some modern parents are becoming aware of the importance of talking with their children about the information presented on television and through other sources. Concerned parents often explain how they perceive the media messages so that their children will learn how to evaluate media images and dialogues. Such intervention helps children to learn how their parents process information from mass media while also passing on the parents' values to the children.

Unfortunately, some children receive no parental advice about the sexual information that is easy to observe in both the electronic and print media. Images and stories of rape and random violence can leave these children adrift with few tools for evaluating media messages about sex. These children are most likely to perceive sexual assault as natural and normal parts of social life, although not all of them will act on that perception.

Girls tend to receive considerably more guidance for evaluating media sex messages than do boys. Many parents and teachers know that naive girls can be seduced or forced into sexual activities that might lead to sexually transmitted infections (STIs), unintended pregnancies, and/or injuries. Thus, many parents, teachers, and media programs tend to target girls more than boys with messages about "responsible sex." Perhaps society should end this double standard and teach more boys about the merits of responsible sex. This can be done in a very positive way by telling boys that they also benefit from reducing the risks of STIs, unintended pregnancies, and violence. In addition, boys could come to realize, "She will love me more if I am responsible, concerned, and loving." Currently, adults may be doing boys a disservice by giving boys fewer guidelines than they give girls for considerate and mature approaches to sex.

2. PUBERTY AND LATER

The interactions of nature and nurture are especially conspicuous during puberty. Because of pubertal hormonal changes, both the sexual organs and secondary sex traits undergo the adolescent changes that make both boys and girls start to look more sexually mature, leading society to increasingly treat them like adults. The hormonal changes also cause the sex organs to produce increasingly pleasurable feelings when touched, speeding the pace of operant learning from sexual experiences.

2.1. Six Basic Sex Differences

During puberty, both boys and girls begin to have increasing numbers of experiences with spontaneous sexual responses and nocturnal orgasms (triggered by unconscious mechanisms deep in the brain), but there are several major sex differences in these hormonally

induced changes. First, pubertal boys begin to have spontaneous sexual arousal 2 to 3 years before girls do, giving boys a significant head start over girls in learning about sex. Second, pubertal sexual arousal is more noticeable and distracting for boys than for girls. A boy who has a spontaneous penile erection is far more likely to notice this conspicuous change than a girl is to notice when she has spontaneous clitoral enlargement and lubrication in the vagina. The girl's response is subtler and more difficult to notice, whereas the boy's response is bigger in that as his penis elongates from 3 or 4 inches (the average length before stimulation) to approximately 6 inches (the average length once erect). Because the boy's penis is well endowed with nerves that connect to the reward centers of his brain, the boy is likely to associate sexual arousal with pleasure. Even though the girl's genitals also connect to the pleasure centers in her brain, the subtly of female sexual responses makes them less noticeable and less likely to create the powerful positive sexual conditioning that boys experience. Third, boys are significantly more likely to experience spontaneous nocturnal orgasms than are girls. Among young people who experience nocturnal orgasms, boys report them approximately four times more frequently than do girls. These orgasms are often associated with sexual dreams, and this strengthens the Pavlovian conditioning that links erotic images with sexual feelings, especially for boys. Fourth, boys experience nocturnal orgasms at a younger age than do girls, giving boys another major head start in sexual learning.

All four of these biological differences lead pubertal boys to become more curious about and interested in sex sooner than do same-age girls, although there is much variability among both boys and girls in the frequency and time of onset of all four of these experiences. Superimposed on these pubertal changes are all of the different levels of inhibitions or positive associations that each individual has learned about sex from his or her parents, family, and culture, with the double standard tending to create more sexual inhibitions for girls than for boys. This interaction of nature and nurture tends to produce the male–female differences in sexual interest shown by the bell-shaped curves illustrated in Fig. 1. The different shapes of the two skewed bell curves show that teen boys tend to be more interested in sex than are same-aged teen girls while also revealing considerable variability in the sexual interest of both boys and girls. For example, some girls fall near the right end of the continuum, indicating that they have more sexual interest than do all of the boys who are to their left on the continuum.

Adolescence brings about two additional concerns about girls' sexuality that can affect their sexual behavior. Sexually active girls tend to worry more about unintended pregnancies—and, in some cases, the issues related to abortion—than do most boys. Also, STIs can cause more serious damage to the reproductive organs of girls than to those of boys; hence, sexually active girls tend to worry more about the risk of contracting STIs than do sexually active boys.

2.2. Double Standards

Societal double standards about sexuality can augment the sex differences seen in Fig. 1. Adolescent girls and adult women are often given the pejorative labels of being "loose" or "sluts" if they have several sexual partners, and this can inhibit some from having too many partners, although there is little consensus on the definition of "too many." In contrast, adolescent boys and adult men often do not feel bad if they have numerous sexual partners and might even be labeled as "studs." Thus, these social labels tend to give girls more inhibitions than boys experience when thinking of having multiple partners. There is another double standard that relates to active and passive social roles. "Nice girls" are not supposed to take active or assertive roles in initiating sexual activities, and many girls feel uncomfortable if they do act too sexually knowledgeable or eager. In contrast, boys are often told that their role is to be active, gain sexual experience, and become sexual experts—and most boys feel happy when they succeed at these. Not surprisingly, in many societies, boys tend to think more about sex and initiate more sexual interactions than do girls.

Studies on the goals for the "first date" among high school and college students reveal that 19- and 20-year-old men typically wanted to touch their partners' genitals (under their clothing) on the first date, whereas same-age women wanted only hugging, physical

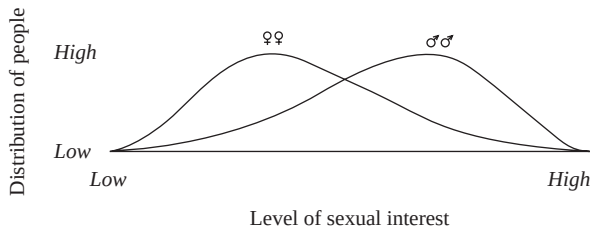


FIGURE 1 Distribution of people along the continuum of sexual interest.

contact, and sensuous kissing. After these students started going out regularly but not exclusively, the men wanted oral sex, whereas the women were just beginning to want their genitals to be touched. When college freshmen were asked whether it was appropriate to have sex if two people had only recently met, 38% of the women and 66% of the men said "yes."

After the "sexual revolution" of the 1960s, many women began to criticize the societal double standards that placed unnecessary inhibitions on female sexuality. As sex education became more widespread, women became increasingly empowered to want to know more about sexuality and have better sex lives. Today, there are increasing numbers of books and magazine articles geared to helping women explore their sexuality, and many more adolescent girls are curious about their sexuality than was true during the 1950s. The development of effective methods of contraception has also allowed both women and men to engage in sex without as much fear of having unintended pregnancies as would have been the case during the 1950s. These social developments have led to more rapid changes in sexual behavior than would be the case if sexual activities were determined by biology alone.

2.3. Recent Trends

Between the onset of the sexual revolution and the early 1990s, there was an increase in adolescent sexual behavior, teen pregnancies, and teen abortions in the United States. However, adolescents have tended to become more sexually conservative since the early 1990s. There has been a decline in the number of teens who have engaged in sexual intercourse, had pregnancies, and obtained abortions. One reason for the reduction in risky teen sexual activities is the societal response to the fear of HIV/AIDS, including increased education about this disease, which has instilled a sense of caution in the age cohorts that reached puberty after 1990.

In the United States, many young adolescents experiment with masturbation, and by 16 years of age approximately 90% of boys and 60% of girls have masturbated. Boys typically masturbate approximately five times per week, and girls do so much less frequently. Typical adolescents begin engaging in sexual intercourse by 16 or 17 years of age, but approximately 20% of students have had sexual intercourse by 15 years of age. During the senior year of high school, 65% of students report having engaged in sexual intercourse. By the end of high school, approximately 20% of adolescents have had four or more sexual partners.

Not all adolescents are certain about their sexual orientation. In a study of junior and senior high school students, 10.7% reported that they were unsure of their sexual orientation and 1.1% reported a bisexual or mainly homosexual orientation. The remaining 88.2% described themselves as predominantly heterosexual. Some of the young people who are uncertain about their sexual orientation experiment with "both sides"; that is, they try sex with both boys and girls so as to understand what their sexual orientation is. Having social support from caring parents, teachers, friends, and counselors can help an adolescent to manage the complex psychological issues that can arise during these years of questioning one's sexual orientation. Before 2003, homosexual behavior was illegal in some states, and this did not help adolescents and young adults to adjust easily to having gay, lesbian, and bisexual identities. In 2003, the U.S. Supreme Court struck down state laws against homosexual behavior, asserting that the constitutional rights to privacy protect people's choices of intimate sexual activities.

Modern society has fewer taboos about sexual behavior than existed a century ago, and more young people are exploring varied sexual activities than in the past. The 100-year trend in sexual interactions shows an increase in petting, oral sex, penile–vaginal intercourse, and anal sex among teens, although since the 1990s young adults have become more cautious about activities they perceive might put them at risk for contracting HIV/AIDS.

3. EARLY SEXUAL PROBLEMS

Some people have an ageist belief that only adults have sexual problems, whereas young people have naturally blissful sex. This is far from true. Although young people can be swept away by the passion and novelty of sex, they often learn about sexual activities through trial and error. Not unexpectedly, many experience numerous errors and frustrations if they have little sex education or prudent guidance in their sexual learning.

The sexual problems that are most common for young adults are understandable in light of the information presented in the prior two sections. Young men tend to become very excited by sexual activities and reach orgasm quickly, which can create the problem labeled as "premature ejaculation" or "early ejaculation." In contrast, young women tend to have limited knowledge about the stimulation they need to reach orgasm, and they have more inhibitions about sexual exploration and playfulness than do young men. As a

result, many young women's biggest sexual problems center on orgasm. For example, they find it difficult to become sexually excited, stay excited, and reach orgasm. Men often reach orgasm too quickly, whereas women tend to have the opposite problem.

The following subsections describe some of the common problems that young people have with sex along with some possible solutions.

3.1. Common Male Problems

Because young men tend to have more experience with spontaneous sexual feelings and masturbation than do young women, young men often know what kinds of stimulation they like best. Many young men discover that using strong and fast sexual movements nearly always leads to orgasm. As a result of this operant learning, young men often develop the habit of masturbating quickly, and this can sometimes bring them to orgasm in a couple of minutes. In fact, rapid masturbation might seem to be a necessity if a young man fears that someone, such as a parent or roommate, may walk in and discover what he is doing if he engages in sexual self-stimulation for a prolonged period of time. Later, when having sex with a partner, the young man might feel the need for rapid sex if he has only the limited privacy offered by a family room or dormitory. All of these early sexual experiences can reinforce and strengthen a young man's habit of using strong and rapid movements to keep sexual activities brief and secret. The fact that his sexual activities at high muscle tension can lead to a very rapid sexual response might not be a problem during masturbation and in some relationships, especially if his partner also fears being caught in the act if the partners take too long.

Although rapid ejaculation might make sense to teens, it can create problems as people mature, have more access to privacy, and begin to explore forms of sexual activity that can extend and heighten the pleasures for both partners. At that time, a man's early habits of rapid movements and early ejaculation can become a liability. This is especially true if the man loses interest in sex after his orgasm but his partner would love to have more sexual interaction. When early ejaculation creates problems in a relationship, it can earn the label of "premature ejaculation." During the 1940s, it was not uncommon for sexual relations to last only 2 minutes before the man ejaculated and sex came to an end. But after the sexual revolution, people have come to realize that it can be pleasurable to extend sexual relations for longer times, especially because some women need 5 or 10 minutes—or longer—to reach orgasm.

There are various methods that men can use to learn how to slow their sexual response and avoid ejaculating too early. Some men can learn the following technique by reading one or two paragraphs on it; however, the stronger a man's habits are for fast and intense self-stimulation, the longer it usually takes to learn to slow down and master new habits for lingering longer over the pleasures of sexual stimulation.

One of the best techniques a man can learn for avoiding early ejaculation is called the "stop–start method." Rather than starting sexual actions and continuing them all the way to orgasm, a man learns to alternate between starting and stopping his physical movements—and his sexual fantasies. Whenever he feels that his sexual activities are causing him to become too excited too quickly, he stops his sexual self-pleasuring and concentrates on bringing pleasure to his partner. Then he resumes cycling back and forth between starting and stopping, learning that the stop periods slow him down and help to ensure that his partner is also experiencing considerable pleasure. Most couples find that the stop–start method brings both of them more gratification than when the man merely starts his most exciting activities and progresses quickly to an early ejaculation.

Men who have learned powerful habits of rapid sex, due to years of positive reinforcement for coming fast, might need to practice the stop–start method systematically for months to unlearn their old habits of speed and discover the multiple rewards of the new habits of self-control. They might need to read Zilbergeld's 1992 book *The New Male Sexuality,* with longer and more detailed descriptions of the technique and several related subtle nuances, to master the new habits.

3.2. Common Female Problems

Women rarely complain about reaching orgasm too fast, although some do. It is much more common for women to report that it is hard for them to become sexually excited, stay excited, and reach orgasm. Among the causes for this are inhibitions about masturbating, talking openly about sex, and seeking sexual information, all of which often result from the double standards of male and female sexual socialization. As a result, many women have a limited understanding of their own bodies, which parts yield the most intense sexual feelings when touched, and how best to stimulate those parts. A lack of knowledge and sexual skill on the part of women—or their partners—can deprive them of the

best types of stimulation for becoming sexually aroused and reaching orgasm.

Most sex therapists and professional self-help books encourage women to experiment with several pleasurable sexual learning experiences. First, women should explore touching all parts of their bodies in the privacy of their bedrooms, perhaps after a soothing bath and with some favorite music in the background. The goal is to discover the parts of their bodies that feel good to touch, noting especially those areas on and around the clitoris and breasts that trigger erotic and sexual feelings. When stimulating those erotic areas, women also need to learn what types of touch feel best, for example, gentle circular motions or back-and-forth stimulation. Some women purchase vibrators to learn how the gentle mechanical oscillations can bring sexual pleasures. Most women learn that gentle stimulation on and around the glans and shaft of the clitoris produce the most sexual excitement, which usually mounts over time, eventually reaching orgasm. Some women discover the effectiveness of water stimulation, either by lying on their backs in a bathtub and allowing water from the spigot to flow over the clitoris or by using a handheld shower to spray water on their external genitals.

Once a woman discovers which parts of her body is erotically sensitive and what types of stimulation she likes, she can progress to either or both of the following techniques. The woman can experiment by herself with self-touching and/or vibrator stimulation to see whether she can reach higher levels of sexual arousal and eventually have orgasms. She can also ask her partner to join her in experimenting with various types of stimulation of her breasts and clitoral area. She and her partner can explore countless different ways of stimulating her clitoral glans and shaft such as gently rubbing up and down, back and forth, or in circular motions. Many women like oral stimulation of the clitoral area, and in fact many more women reach orgasm through oral stimulation than by penile–vaginal intercourse. Both women and men need to realize that stimulation of the inside of the vagina with a penis, dildo, or penis-shaped vibrator is far less likely to lead to female orgasm than is stimulation around the clitoral area. Because men usually think of the penis as their key sex tool and source of pleasure, they often orient their sexual activities to penile–vaginal intercourse, but this can easily deprive women of stimulation to the clitoral glans and shaft.

To create additional sources of erotic stimulation that can augment the physical arousal derived from stimulation of the breasts and clitoral area, sex therapists often encourage women to develop sexual fantasies. For centuries, men have paired erotic images and fantasies with the sexual excitement of masturbation and coitus, thereby turning their fantasies into sexual turn-ons through Pavlovian conditioning. Increasing numbers of women are learning to eroticize mental images and fantasies by pairing them with the sexual excitement of masturbation and partner stimulation.

Another problem that females report more often than do males is painful intercourse, which is technically called "dyspareunia." Anyone who experiences painful intercourse should see a physician for a personalized diagnosis and medical advice because there are many possible causes of sexual pains. In serious cases, it is prudent to avoid penile–vaginal sex until after consulting with a physician. People who cannot abstain from penile–vaginal intercourse might try using a water-based lubricant such as KY Jelly or Astroglide to reduce the friction of coitus. For women with dyspareunia, manual or oral stimulation to the clitoris and surrounding areas can lead to lubrication and orgasm before penile intromission—and even without it. In cases of female pain, another solution involves a much slower and more gradual introduction to erotic play and intromission than most males tend to do without a special request. If these practices are not helpful, a physician should be able to identify the exact cause of the pain and suggest at least one solution. For example, the woman may have a vaginal infection or a urinary tract infection. Either the woman or her partner may be allergic to the latex in condoms. A number of useful Web sites provide additional information about both male and female sexual problems.

During the past half-century or so, there have been so many advances in the science of sexuality that increasing numbers of people are enjoying the pleasures of sex without the problems of pain, unintended pregnancies, or STIs. Unfortunately, society still has a long way to go in reducing the persistent problems of sexual assault, rape, domestic violence, HIV/AIDS, and homophobia. The fact that sexuality can be discussed, written about, and legislated more freely today than was the case a mere 50 years ago bodes well for continued progress in resolving these remaining problems.

See Also the Following Articles

Gender and Culture ■ Homophobia ■ Homosexuality ■ Risk-Taking and Sexual Behavior ■ Sexual Behavior and Culture

Further Reading

Baldwin, J. D., & Baldwin, J. I. (1981). *Beyond sociobiology*. New York: Elsevier.

Baldwin, J. D., & Baldwin, J. I. (1997). Gender differences in sexual interest. *Archives of Sexual Behavior*, *26*, 181–210.

Friedrich, W. N., Fisher, J., Broughton, D., Houston, M., & Shafran, C. R. (1998). Normative sexual behavior in children: A contemporary sample. *Pediatrics*, *101*, E9.

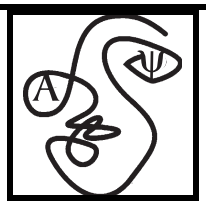
Kaeser, F., DiSalvo, C., & Moglia, R. (2001). Sexual behaviors of young children that occur in schools. *Journal of Sex Education and Therapy*, *25*, 277–285.

Pfaus, J. G., Kippin, T. E., & Centeno, S. (2001). Conditioning and sexual behavior: A review. *Hormones & Behavior*, *40*, 291–321.

Remafedi, G., Resnick, M., Blum, R., & Harris, L. (1992). Demography of sexual orientation in adolescents. *Pediatrics*, *89*, 714–721.

Sargent, T. O. (1988). Fetishism. In D. M. Dailey (Ed.), *The sexually unusual: Guide to understanding and helping* (pp. 27–42). New York: Harrington Park Press.

Zilbergeld, B. (1992). *The new male sexuality*. New York: Bantam Books.



Sexual Behavior and Culture

Lise D. Martel, Skyler Hawk, and Elaine Hatfield

University of Hawaii at Manoa, Honolulu, Hawaii, USA

1. Sexual Permissiveness
2. Sexual Expression
3. Contraception and Birth Control
4. Sexuality and HIV/AIDS
5. Conclusions
   Further Reading

GLOSSARY

agrarian Living from the management or distribution of land.
AIDS Abbreviation for acquired immune deficiency syndrome.
double standard Arbitrary use of two different sets of rules for two distinct groups.
globalization The sharing and adoption of new techniques resulting from the increasing interconnectedness of the nations of the world.
HIV Abbreviation for human immunodeficiency virus, which is responsible for the development of AIDS.
hunter–gatherers Nomadic people who live by hunting wild game, fishing, and gathering nuts, fruits, berries, and other plant products growing in the wild.
industrial society Community where industry is the principal activity of the society.
pandemic Disease affecting a whole country or the whole world.
sexual permissiveness Toleration of a wide variety of sexual behaviors and attitudes.

Sexual behaviors and values are important components of all cultures. Societies may differ greatly in the meanings they attribute to sex, their sexual permissiveness, the expressions of their sexual values, their views on contraception and pregnancy control, and the health issues they face. Yet despite their infinite variety and the individual differences within each culture, some broad generalities can be identified. Societies' commonalities and differences in behaviors and beliefs are determined by cultural, socioeconomic, and historical contexts, and during recent years they have been determined by globalization and technological advancement as well. By engaging in social comparison, people garner their understandings of naturalness and normality, thereby forming their attitudes toward sexual issues and determining what is sexual and sexually acceptable.

1. SEXUAL PERMISSIVENESS

A society's general level of sexual permissiveness most likely originates from complex interactions among several cultural factors. For example, a culture's technological sophistication, the influence of particular faiths, and the importance of economic exchange at marriage all may have impacts on the place of sex in a society. Acceptable forms of sexual expression are also determined by the importance that is placed on reproduction. For example, engaging in homosexual activity may be viewed more negatively in those cultures that hold procreation, but not pleasure, as an acceptable motivation for sex.

©2004 Elsevier Inc.
All rights reserved.

1.1. Permissive Societies

Many cultures see sexual expression as natural and beneficial. Cultures that can be labeled as sexually permissive hold generally positive attitudes toward sexual development, which is seen as essential to healthy social adjustment. There are usually fewer social restrictions placed on sexuality. Children in these societies are not sheltered from sexual knowledge. As with the Muria of central India, they might even be highly exposed to nudity and sexual activity through ritual and institutional practices. This is not seen as problematic because healthy sexuality is considered to be necessary for marital happiness and mature functioning. Both men and women in such cultures tend to equally value pleasure associated with sexual expression. Premarital sex is commonly accepted as a reality for both men and women.

1.2. Restrictive Societies

In contrast, sexually restrictive cultures attempt to limit both the occurrence of sexual expression and the ways in which it can be expressed. Sex itself is either seen as holding a unique and dangerous power or considered to be harmful because it is closely connected with negative outcomes such as early pregnancy and venereal disease transmission. Adults in restrictive societies generally try to limit children's understanding of sexuality for longer periods of time, but for different reasons. For example, in the United States, parents' long-term goals for children's development include later marriage and prolonged education. Because the possible negative consequences of sexual activity are seen as antithetical to these goals, abstinence may be emphasized and may become the cultural norm for moral and acceptable behavior. Thus, in restrictive cultures, virtue is often associated with premarital chastity. Sexual abstinence before marriage is stressed for one sex (usually women) or both sexes. In cases where women's chastity is emphasized, men are typically permitted at least some greater amount of sexual freedom.

1.3. Sexual Double Standard

When inequalities in sexual freedom are found, sexual activity and associated pleasure tend to be more valued by the gender whose behavior is less restricted. This describes what is referred to in Western literature as the sexual "double standard." In many cultures, men are often encouraged to seek out and acquire sexual experiences with many partners, whereas women are expected to delay sexual activity until after marriage. This double standard can produce different sexual motivations for men and women. Men are more likely to endorse reasons of enjoyment and pleasure, whereas women are more likely to endorse reasons of love and emotional commitment. Although this double standard has eroded considerably in industrial countries, it can still be observed under some circumstances in that people are generally more tolerant of men who engage in casual sex or unusual sex practices than they are of women who do so. In most African and Asian nations, a double standard is still frequently reported.

2. SEXUAL EXPRESSION

Sexual expression varies widely from one culture to another. These differences include, but are not limited to, cultural views and behaviors related to premarital and extramarital sexuality, homosexuality, transgenderism, genital mutilation rituals, contraception and birth control, and issues related to the HIV/AIDS pandemic.

2.1. Premarital Sexuality

2.1.1. Agrarian and Hunter–Gatherer Societies

In agrarian and hunter–gatherer societies, premarital intercourse is often perceived negatively. Although the absence of reliable birth control and the disastrous social consequences of unwanted pregnancies are good incentives for compliance, early age of marriage and late sexual maturation also make this social expectation bearable. In such cultures, a promise of marriage is often the only acceptable situation where premarital intercourse is considered acceptable. For some cultures, it is sexual maturity rather than marriage that counts. For the !Kung of southern Africa, married men may take prepubescent brides but must wait for their menarches to have sex with them.

2.1.2. Industrial Societies

In many industrial societies, the increasingly permissive attitudes toward premarital sexual activity can be partly explained by women's empowerment through oral contraception and advances toward economic equality between men and women. Also, delaying marriage until a later age increases the gap between sexual maturity and marriage and makes premarital abstinence an unrealistic

goal for many. Yet despite the prevalence of more permissive attitudes, for most young people in industrial societies, sex is still expected to involve some level of affection and commitment between partners.

2.2. Extramarital Sexuality

2.2.1. Agrarian and Hunter–Gatherer Societies

In many agrarian and hunter–gatherer societies, extramarital relations are considered to be "appropriate" for men under certain socially established conditions. For example, in Thailand, young men are viewed as having insatiable sexual drives that they have a right to satisfy, even after marriage, through visits to sex workers. In some African societies, the importance of a large family to a man's status gives him the right to copulate with another woman if his own wife is infertile.

2.2.2. Industrial Societies

Researchers agree that two changes are evident during recent years in industrial societies. First, the sexual double standard is eroding. Men and women are becoming similar in their willingness to experiment with extramarital sex. Second, men and women are experimenting with extramarital relations earlier, more often, and with more partners than was the case in the past.

2.3. Homosexual Orientation and Behaviors

Homosexuality generally refers to sexual attraction or behaviors toward people of one's own biological sex. Differences in how "sex" is defined in a particular culture, and the meanings that are imparted on different sexual behaviors between people of the same sex, can often lead to contrasting ideas about what "homosexuality" actually entails and what behaviors are associated with it. In cultures where sex is defined in very narrow terms, such as consisting solely of penile–vaginal intercourse, a great deal of intimate behavior may be acceptable between people of the same sex. Sometimes, behaviors that would be considered highly sexual in one society are given a different nonsexual meaning in other cultures. For example, young adolescent boys in the Sambia culture of Papua, New Guinea, ritually fellate older men as a way of absorbing their strength and masculinity. However, they are later expected to continue on to exclusive heterosexual marriages and relationships, and adult same-sex relationships are not well accepted. In Ancient Greece, important older men often took younger men as sexual partners concurrently with the older men's heterosexual marriages.

Homosexuality is highly connected to perceptions of gender roles in a number of cultures. For many peoples in Latin, African, and Middle Eastern countries, as well as among many ethnic minorities in America and Europe, men may engage in same-sex behaviors yet still identify as "heterosexual." They are able to engage in these behaviors, without incurring the stigma that accompanies a label of homosexuality, due to the roles that they play in these sexual activities. Men in these cultures who take the "inserting" role in same-sex encounters are still seen as masculine and heterosexual, presumably because they are adhering to sexual–cultural norms for their gender. On the other hand, men who take the "insertee" role, which is commonly viewed as the woman's role in sex, are typically viewed as feminine and are labeled "homosexual."

In contrast, in the United States and much of Western Europe, homosexuality is perceived to be an intrinsic aspect of the individual rather than being defined by gender roles. Therefore, men who identify as homosexual experience a much greater interchangeability in their sexual role taking. The Western conceptualization of homosexuality has been bolstered during recent years by a plethora of scientific information and political assertions that homosexual orientation has a genetic etiology.

As globalization continues to proliferate, many cultures are shifting their notions of homosexual behavior to be more consistent with Western (particularly American) ideologies. Still, cultural differences are numerous. In multicultural societies such as the United States, cultural differences in the conceptualization of homosexuality should be considered in social marketing campaigns, as well as in other interventions, that are geared toward homosexuals of specific minority groups.

2.4. Transgenderism

The term "transgender" applies to people who experience a deep lifelong preference for the identities or gender roles usually associated with being a member of the opposite sex. Some people who consider themselves to be transgendered choose to live out the conventional roles assigned to them by society. Many others take a variety of approaches to socially identify with what they feel is their true gender. Some choose to wear clothes and ornaments of the opposite sex; adopt the postures, gestures, behaviors, and roles of the

opposite sex; and/or undergo sex reassignment surgery. They might become transvestites or cross-dressers, drag queen or kings, or transsexuals. Yet this does not imply that becoming a cross-dresser or a drag queen or king is always motivated by transgenderism.

In most Western societies, the challenges that transgenderism poses to traditional Western gender roles still result in widespread stigmatization. In fact, in Western culture, transgenderism, transvestism, and transsexuality are often labeled as forms of a mental disorder called gender dysphoria. In contrast, in many non-Western societies, transgendered individuals, transvestites, and transsexuals are valued for their uniqueness. For example, historically in the Zuni culture of New Mexico as well as in many island cultures of the Pacific, male children who displayed feminine characteristics were considered to be endowed with great power. They dressed as a woman and performed functions of healer and arbiter. In today's Samoa, humorous entertainment provided by cross-dressers is used to diffuse violence in public gatherings. In India, the Hijras, a group of men who have had their genitalia removed and have sex only with other men, make a living by dancing, singing, and offering blessings at weddings. During recent years, many Hijras have been elected to public office due to a perception that they are less likely to be corrupted than are other public officials.

2.5. Genital Modification Rituals

Male circumcision is probably the most common form of genital modification. From Ancient Egypt to modern-day Africa and America, circumcision has been practiced under all kinds of justifications. In America, circumcision was popularized into the 1900s as a cure for masturbation and as a way in which to ensure good hygiene and prevent diseases. Today, undergoing circumcision is still a common practice in many cultures, especially Jewish and Moslem cultures.

In many Sub-Saharan societies and Moslem Arab groups, women undergo genital alterations and mutilations. The modifications range from removing the clitoral prepuce (sunna circumcision) to removing the entire clitoris and most of the labia and stitching together the two sides of the vulva, allowing only a small opening for the passage of urine and menstrual blood (infibulation). Many of these gruesome procedures result in health problems or even death (in extreme cases). In the United States, they have been outlawed since 1996.

2.6. Future Directions

The meanings that a culture gives to sex, the sexual expressions it recognizes, and its level of permissiveness are in constant flux. Globalization and rapid technological development are increasing the rate of such changes and continually force all cultures to reexamine their beliefs about issues related to sexual permissiveness and sexual expression. Many Western practices, such as prolonged adolescence and delayed marriage, serve to remove sexuality from family control and place it in the hands of the individual. Agrarian and hunter–gatherer societies increasingly are adopting the ways of the West, forcing some rapid changes on their cultures. Some changes, such as relaxing premarital sexual standards and promoting gender equality, permit more individual freedom. Other changes, such as laws regulating genital modification rituals, impose new restrictions on cultures.

3. CONTRACEPTION AND BIRTH CONTROL

In both developing and industrial societies around the world, women are perceived to be responsible for preventing unwanted pregnancies. In many developing nations, women are under the economic control of men, have little access to modern contraceptives, and are expected to bear the responsibility of preventing unwanted pregnancies even if they have low control. Not surprisingly, most of the world's births occur in developing nations. In industrial societies, the situation of women is quite different. Success of pregnancy prevention is increased by gains in equality of rights, economic independence, and access to modern contraceptives that do not require the cooperation of their male partners.

3.1. Developing Societies

In rural cultures of the developing world, children are economic assets. From an early age, they become active contributors to the survival of their families through a variety of labor tasks such as those related to food cultivation and sibling caretaking. In an effort to maximize efficiency of the family unit and reduce health risks, women in these societies generally try to maintain a birth spacing of 2 to 3 years. To control this spacing, they usually rely on abstinence and on the reduced fertility offered by lactation. In many African and Southeast Asian cultures, there is an expectation that

husbands will abstain from having sex with their wives for a period of time after their wives give birth, but husbands are not necessarily expected to cease sexual activity completely. Extramarital sex, visits to sex workers, and (to a lesser extent) polygamy are common culturally acceptable alternatives for men. Women are usually expected to remain abstinent postpartum.

Despite the fact that a large number of people in developing countries live in rural areas, as a result of globalization and industrialization, an increasingly larger number live in overcrowded cities. For these people, the economic assets provided by children are often replaced by the children becoming an additional strain on already compromised financial means.

3.2. Industrial Societies

In most industrial societies, where childhood is usually spent in school, children represent an economic cost rather than an asset. Pregnancy is often planned and heavily controlled by the individual through the use of contraceptives. In some countries, such as China, pregnancy is also restricted by the state in an effort to manage overpopulation.

In the United States, at the time of this writing, the current political administration promotes abstinence as the only acceptable alternative to prevent pregnancy and sexually transmitted diseases among teenagers. Yet many organizations have successfully promoted the use of various modern forms of contraception to prevent pregnancy in teenagers and adults alike. Recent trends show a decline in unwanted pregnancies among U.S. adolescents. Still, the United States maintains a disproportionate number of abortions in comparison with other Western industrial countries. In countries such as Holland, Sweden, France, and Great Britain, it is believed that the general endorsement by individuals, society, government officials, and the media of the message that protected sex (rather than abstinence) is expected from teenagers has facilitated the decline in abortions and the delay in the onset of sexual activity.

3.3. Future Directions

Under the influence of globalization, beliefs about individualism, human rights, and gender equality are having a growing impact on the sexual behaviors and attitudes of people from developing nations. Unfortunately, these nations usually lack the medical and financial resources to adopt modern contraceptive methods that have proven to be successful in the wealthier industrialized nations. In the increasingly interconnected world, overpopulation in developing countries is a universal concern. All societies must collaborate to address the old and new challenges that arise from living in an interconnected world.

4. SEXUALITY AND HIV/AIDS

The awareness of the link between sexuality and health is not new. An estimated 333 million new cases of curable sexually transmitted diseases occur worldwide every year. Yet it was only after the first diagnosed case of AIDS that sexually transmitted diseases came to the forefront of global health concerns. Since the early 1980s, more than 42 million people worldwide have been infected with HIV, making HIV/AIDS the most devastating sexually transmitted disease in the world's history.

4.1. Condom Use

4.1.1. Developing Societies

HIV/AIDS has been especially devastating in Sub-Saharan Africa, Southeast Asia, and Latin America. In an effort to control the spread of HIV/AIDS, the most common approach used worldwide has been to promote the use of male condoms and provide education on the virus and its transmission. This approach is especially adequate in developing countries where most of the transmission results from heterosexual intercourse. Yet the success of this intervention varies widely across cultures and even between genders. Many societies do not have the literacy, education, and access to condoms that are necessary for successful prevention of HIV/AIDS. The use of condoms is often heavily stigmatized. Beliefs that condoms are appropriate only with prostitutes and that only promiscuous women carry condoms are common. Health beliefs can also interfere with the use of condoms. For example, in Cambodia, condoms are believed to disrupt the balance of the body and burn the womb.

Nonetheless, condoms have had a tremendous impact on the AIDS epidemic in many developing countries. For example, in Southeast Asia, where the commercial sex industry had fueled the AIDS epidemic, several countries have succeeded in increasing condom use among sex workers and their clients. Thailand, with its 100% condom use policy in brothels, is a prime example of the successful implementation of a condom use program in a high-risk population.

4.1.2. Industrial Societies

In industrial countries, the male condom is still the only effective protector against HIV/AIDS. During recent years, many attempts have been made to develop competing alternatives. The male condom has been criticized for being a barrier to women's empowerment in protecting their health. It was hoped that the female condom, which consists of a malleable tube with a flexible ring at both ends (one that can be inserted into the vagina by the woman the other resting at the opening of the vagina), would give women the empowerment they had been seeking. Unfortunately, proper use of the female condom still requires active cooperation from their male partners during penile insertion. In addition, despite promising theoretical efficacy, its use efficacy proved to be much lower than that of the male condom.

4.2. Gender Differences

4.2.1. Developing Societies

In developing societies, HIV/AIDS is equally prevalent among men and women. Yet women are at greater risk for acquiring HIV than are men because women usually are less educated and do not have the societal position or financial means to obtain condoms. Their disadvantage is further compounded by the fact that HIV is more easily transmitted from men to women. Furthermore, the only known effective barrier to HIV infection, condoms, requires male cooperation. For women, risk of infection is sometimes exacerbated by the practice of dangerous cultural habits. For example, in Sub-Saharan Africa, which has the highest prevalence of HIV/AIDS in the world, the custom of dry sex is common. In this practice, women prepare themselves to pleasure their husbands by drying their vaginas with a mixture of powdered stem and leaf and water, soil mixed with baboon urine, detergents, salt, cotton, or shredded newspaper. These increase dryness, friction, vaginal lacerations, and the suppression of the their vaginas' natural bacteria that, in turn, increase women's risks of acquiring sexually transmitted diseases, especially HIV/AIDS. The Muslim Hausa of Nigeria perform cuts on the walls on the vagina to cure women suffering from a variety of maladies. In addition to exposing the women to risk of hemorrhage and infection, the cuts provide the AIDS virus with direct access to the bloodstream during intercourse.

Targeting women with interventions in developing countries could have a large impact on the HIV/AIDS pandemic because women can also transmit the HIV infection to their infants through delivery and breast-feeding. In developing countries, women rarely have access to highly active antiretroviral treatments to limit the transmission during delivery. Also, in countries with high prevalences of waterborne diseases, such as cholera, breast-feeding is often the only viable means of survival for infants. Therefore, prevention interventions with mothers would also have a significant impact on the health of their children.

4.2.2. Industrial Societies

In Western countries, since the 1980s, gay or bisexual men and intravenous drug users have been recognized as high-risk populations for HIV/AIDS. Still today, 82% of the estimated cumulative 877,275 cases of AIDS in the United States are among men. Of these, 55% are those who have sex with other men and 26% are injection drug users. Although women of childbearing age are the fastest growing population of AIDS cases in America, 70% of new HIV infections still occur among men. Yet because of prevention intervention efforts, the estimated number of new HIV infections in the United States has decreased from approximately 100,000 to 40,000 per year. Moreover, with the advances in HIV treatments, especially the advent of highly active antiretroviral treatments, great progress has been made in reducing the number of new AIDS cases as well.

4.3. Future Directions

Efforts should continue to be made to improve women's control over their own protection against HIV/AIDS. More likely, the solution will come in the form of a microbicide if issues of genital irritation can be resolved.

In both developing and industrial countries, one group that has not been adequately targeted in HIV/AIDS intervention is adolescents. In addition to being behaviorally and physiologically more vulnerable to acquiring HIV/AIDS and other sexually transmitted diseases, adolescents are often socially disadvantaged. They lack independent incomes and often are not offered the privacy of information that is necessary for open communication with health care providers.

Currently in the United States, funding for prevention efforts is shifting from targeting noninfected individuals to targeting infected individuals. The progress made in preventing new cases of HIV/AIDS seems to justify this new approach, yet efforts should be

maintained to continue prevention efforts with high-risk populations and fast-growing groups.

5. CONCLUSIONS

Globalization has propelled the societies of the world in an interaction that transforms cultures and redefines the sexuality of every nation. It has both positive and negative effects. On the one hand, it encourages the sharing of perspectives and medical advancements that can help societies to adopt sexual norms and beliefs that are compatible with healthy lives. On the other hand, it facilitates the spread of disease and reduces cultural diversity in the human sexual experience. It is hoped that in the future, through collaboration between industrial and developing nations, both healthy sexuality and gender equality will become a part of every distinct society and discrimination will become a rarity in the lives of individuals who belong to stigmatized groups such as homosexuals, transgendered individuals, and those infected with HIV/AIDS.

See Also the Following Articles

Gender and Culture ■ Homosexuality ■ Marital Therapy ■ Risk-Taking and Sexual Behavior ■ Sexual Behavior

Further Reading

Ford, C. S., & Beach, F. A. (1951). *Patterns of sexual behavior*. New York: Harper & Row.

Francoeur, R. T. (Ed.). (1997). *The international encyclopedia of sexuality* (Vols. 1–3). New York: Continuum.

Francoeur, R. T., & Noonan, R. J. (Eds.). (2002). *The international encyclopedia of sexuality* (Vol. 4). New York: Continuum.

Galliano, G. (2003). *Gender: Crossing boundaries*. Belmont, CA: Thomson Wadsworth.

Gibney, L., DiClemente, R. J., & Vermund, S. H. (Eds.). (1999). *Preventing HIV in developing countries: Biomedical and behavioral approaches*. New York: Kluwer Academic/Plenum.

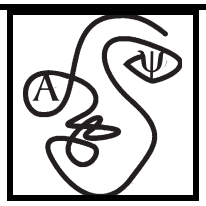
Hatfield, E., & Rapson, R. L. (1993). *Love, sex, and intimacy: Their psychology, biology, and history*. New York: HarperCollins.

Hatfield, E., & Rapson, R. L. (1996). *Love and sex: Cross-cultural perspectives*. Needham Heights, MA: Allyn & Bacon.

Hotvedt, M. E. (1990). Emerging and submerging adolescent sexuality: Culture and sexual orientation. In J. Bancroft, & J. M. Reinisch (Eds.), *Adolescence and puberty: The Kinsey Institute series* (Vol. 3, pp. 157–172). New York: Oxford University Press.

LeVine, R. A., Dixon, S., LeVine, S., Richman, A., Leiderman, P. H., Keefer, C. H., & Brazelton, T. B. (1998). *Child care and culture: Lessons from Africa*. Cambridge, UK: Cambridge University Press.

Shostak, M. (2001). *Nisa: The life and words of a !Kung woman*. Cambridge, MA: Harvard University Press.



Sexual Harassment

Janet Sigal
Fairleigh Dickinson University, Teaneck, New Jersey, USA

1. Prominent Cases and Laws Against Sexual Harassment
2. Theories and Models of Sexual Harassment
3. Measures of Sexual Harassment
4. Incidence of Sexual Harassment
5. Outcomes and Responses
6. Prevention and Intervention
7. Conclusion
Further Reading

GLOSSARY

gender harassment Taunts, gender-based insults and gestures, and/or the presence of pornographic materials in the workplace.

hostile environment sexual harassment When working conditions serve as a source of distress for the employee.

quid pro quo sexual harassment Requesting sexual favors in exchange for promotion or other employment-related decisions.

sexual coercion Job-related threats and/or bribes to victims for giving sexual favors.

sexual harassment Unwelcome sexual advances, requests for sexual favors, and other verbal or physical contact of a sexual nature when (a) submission to such conduct is made, either explicitly or implicitly, a term or condition of an individual's employment; (b) submission to or rejection of such conduct by an individual is used as the basis for employment decisions affecting the individual; or (c) such conduct has the purpose or effect of substantially interfering with a person's work performance or creating an intimidating, hostile, or offensive working environment.

unwanted sexual attention Persistent requests for dates (without any sanctions or job-related threats associated with refusals), unwanted touching, and/or harassing victims through phone calls or letters.

Sexual harassment is a pervasive and persistent problem in the United States as well as around the world. According to estimates, approximately half of the women in the United States will experience sexual harassment at some time, either in the workplace or in educational settings. Men also are victims of sexual harassment, although at a lower rate. Sexual harassment occurs internationally at a rate similar to, or possibly even higher than, that in the United States. Because this problem affects many people worldwide, it is an issue that must be explored and investigated. Before prevention programs can be developed, laws must be understood, the incidence of sexual harassment worldwide must be examined, the consequences for victims must be established, and models must be explored to determine why sexual harassment occurs.

1. PROMINENT CASES AND LAWS AGAINST SEXUAL HARASSMENT

1.1. Prominent Cases

Sexual harassment became the focus of intensive research during the 1990s, due partly to the existence of

©2004 Elsevier Inc.
All rights reserved.

high-profile cases in the United States. In 1991, during a televised U.S. Senate hearing in which Clarence Thomas was being considered for appointment to the U.S. Supreme Court, professor Anita Hill accused Thomas of having sexually harassed her when she had worked for him at the Equal Employment Opportunity Commission (EEOC). Although Thomas was confirmed as a Supreme Court justice, the widely viewed hearings brought the issue of sexual harassment to the attention of the American public. Another highly publicized 1991 event, in which several naval officers were accused of sexually harassing women at the Tailhook convention for naval aviators, also was responsible for focusing public attention on this issue. Whether sexual harassment occurs in the workplace (e.g., when an employer demands that an employee exchange sexual favors for a promotion) or in educational settings (e.g., when a professor demands sexual compliance for a high course grade), the behavior is illegal and poisons the climate of the workplace or educational institution.

1.2. U.S. Sexual Harassment Laws

Table I lists some cases and the precedents set by these court decisions. Sexual harassment laws related to the workplace stem from Title VII of the Civil Rights Act of 1964 (and the extension of this act in 1991) that prohibits any form of discrimination. Laws against sexual harassment have been developed within the framework of sex discrimination in that harassment targets employees due to their gender and prevents these individuals from obtaining equal opportunities in the workplace. Laws prohibiting academic sexual harassment are related to Title IX of the Education Amendments.

In 1987, a court supported the establishment of the reasonable woman standard concept. This concept developed as a consequence of studies demonstrating that women were more likely to perceive inappropriate sexually related behaviors as sexual harassment and to expect organizations to rectify the situation than were men. Because women tend to be at higher risk for sexual harassment than are men, women might feel more threatened and more sensitive to the presence of harassment behaviors constituting a hostile environment. Therefore, it was reasoned that if jurors were encouraged to determine whether a reasonable woman would be upset and made to feel uncomfortable by these behaviors, they might be less biased against a woman who alleges that she has been sexually harassed than if the jurors adopted a reasonable person standard. It has not yet been clearly demonstrated that the reasonable woman standard is superior to the reasonable person standard in court cases.

TABLE I
Significant Sexual Harassment Court Decisions in the United States

Court case	*Decision*
Davis v. County Board of Education (1999)	Schools will be held responsible for peer harassment if it is severe and the schools were aware of the harassment.
Oncale v. Sundowner Offshore Services (1998)	Same-sex harassment and harassment of men constitute actionable sexual harassment.
Lois E. Jensen v. Eveleth Taconite Co. (1997)	First class-action sexual harassment hostile environment lawsuit was filed.
Harris v. Forklift Systems (1993)	"Abusive and hostile workplace" was established as a precedent for lawsuits, whether or not the employee suffered "psychological damages."
Yates v. Avco Corporation (1987)	The "reasonable woman standard" was established.
Meritor Savings Bank v. Vinson (1986)	A clearer definition of "hostile environment" was provided, indicating that it involves "unwelcome behavior;" stating that the behavior had to be "persistent and severe" for it to be considered "actionable;" and limiting the employer's liability to when the incidents were "known or should have been known."

1.3. International Laws Against Sexual Harassment

Because the world is globally connected, it is important to examine international laws against sexual harassment. In 1992, a study indicated that only 9 of 23 countries examined had specific laws against sexual harassment. In 2000, 31 countries that had legal statutes against sexual harassment were identified, but it

was suggested that punishment and liability for the behavior vary from country to country due to different legal systems and conceptualizations of the issue.

Legislation in Canada and the European Union (EU) is very similar to that in the United States. An interesting comparison can be made between the United States and France, the model for the EU legislation, in terms of laws relating to sexual harassment (or *harcelement sexuel*). Quid pro quo and hostile environment sexual harassment laws stemming from a sex discrimination framework have been the basis of lawsuits in the United States, and plaintiffs have received large monetary compensation in civil trials. For example, a $7.1 million judgment was assessed against a major law firm and one of its former partners. French laws, however, focus entirely on quid pro quo sexual harassment emanating from an authority figure as the basis for allegations of sexual harassment that is considered to be a form of violence against women. Because sexual harassment is part of the French Penal Code, presumably a transgressor may be sent to prison for this offense. However, that consequence is unlikely to occur, and the maximum fine imposed ($16,000) is far less than that imposed by U.S. juries in sexual harassment cases. French law also does not cover peer harassment. Legislators used anti-American sentiment to denounce initial attempts to develop French sexual harassment laws as "importing American excesses of litigiousness, Puritanism, and the battle of the sexes."

In the Far East, legislation to prevent sexual harassment has been passed in Australia, Bangladesh, Japan, The Philippines, Sri Lanka, and Hong Kong. Court guidelines against sexual harassment exist in India as well. Although Mainland China has amendments, such as the "Law on Safeguarding the Rights of Women," these guidelines have been criticized due to a lack of enforcement and penalties.

In Latin America, few countries have well-developed laws prohibiting sexual harassment. This is partly because harassing behaviors are not perceived as unwanted in some of these countries.

2. THEORIES AND MODELS OF SEXUAL HARASSMENT

There are three major categories of models: power models, biological and evolutionary models, and an integrative organizational model.

2.1. Power Models

Several approaches suggest that power is the dominant force behind sexually harassing behaviors. The organizational model states that because men are disproportionately represented in powerful positions of organizational authority, they use these positions to gain sexual satisfaction. The sociocultural model suggests that men have been socialized and rewarded for being aggressive and dominant and that women have been socialized for being submissive and nonassertive. These rewarded sex role behaviors then transfer into the workplace setting and serve to discourage women from competing in the "man's" world of work. As an example, one male employee stated that it took 6 months before a woman coworker in his department "got the message that she wasn't welcome and didn't belong." The sex role spillover model proposes that when women are in nontraditional workplaces, they are responded to in terms of gender identity instead of role identity; that is, they are treated differently from male coworkers and are treated as women as opposed to, for example, police officers or construction workers.

Although these power models have received some research support, they appear to be too simplistic to account for all sexual harassment, including peer harassment when there are no power differentials. In addition, these approaches fail to explain why some men exploit their positions and engage in sexual harassment and others do not.

2.2. Biological and Evolutionary Models

The biological and evolutionary models suggest that aggressive sexual behavior is natural and expected for men, both in the workplace and elsewhere. Research generally has not supported this approach.

2.3. Integrative Organizational Model

An integrative organizational model proposes two essential conditions that are favorable toward the expression of sexual harassment behaviors: climate and gender ratio in the workplace.

2.3.1. Climate

This factor refers to whether or not the organization has a climate that is supportive of sexually harassing behaviors. Although there may be an organizational policy against sexual harassment, if victims are afraid to come forward with complaints due to negative consequences, this climate would be classified as supportive of sexual harassment. For example, a female police officer who reported being sexually harassed was told by her supervisor that she could go forward with her complaint but that if she did, she could not expect her backup to appear when she called even if she was in a dangerous situation in the field and her life was on the line.

2.3.2. Gender Ratio in the Workplace

Sexual harassment is more likely to occur in a workplace where there are more men than women and where women are performing in nontraditional male-oriented jobs. Law enforcement and the military in the public sector, as well as private sector occupations such as construction work, might be related to this type of category.

The model suggests that sexual harassment has certain consequences for the victim, including lowered job satisfaction that can be associated with work withdrawal or work avoidance (e.g., frequently arriving late for work), health outcomes (e.g., headaches) that can be connected to job withdrawal (e.g., quitting the workplace), and psychological outcomes (e.g., anxiety, depression). The model proposes the following factors that help to determine the employee's reaction to sexually harassing experiences:

- Personal vulnerability (e.g., younger women are more susceptible to harassment)
- Response style (i.e., how the victim attempts to cope with the harassment experience)

Although the integrative model has been supported by U.S. and cross-cultural research, it also has been criticized for not indicating how sexual harassment occurs in the first place and why some men engage in sexual harassment and others do not. In addition, it does not specify the extent to which subjective appraisals, affected by the severity and frequency of the experiences, contribute to the effects of the harassment or the way in which additional job stresses (e.g., racism in the workplace) may interact with sexual harassment.

3. MEASURES OF SEXUAL HARASSMENT

A number of methodological problems characterized earlier incidence surveys that produced widely varying estimates of harassment behaviors. For example, two people could accurately state on a general measure that they had been sexually harassed, although the severity of the harassment incidents may have varied considerably. Another problem related to retrospective accounts (e.g., surveys that ask how often the respondent has been sexually harassed during the past 5 years) that are subject to memory issues as well as changes in social norms. Other issues involve the lack of psychometric data for the various scales and response biases in return rate that often limit the usefulness of these surveys.

Currently, the primary measure of sexual harassment in the United States, as well as in some other countries, is the Revised Sexual Experiences Questionnaire (SEQ-R), an 18-item self-report measure developed by Fitzgerald and colleagues in 1995. The SEQ-R is the most reliable and valid measure available and has been used in many settings, including workplaces, universities, and the military. It describes an extensive listing of behaviors that constitute gender harassment, unwanted sexual attention, and sexual coercion. Most important, it does not ask respondents whether they have been sexually harassed until the end of the questionnaire. Studies have shown that respondents often answer that they have not been sexually harassed, even if they have experienced several types of harassment behaviors, possibly because participants have different definitions of sexual harassment, are reluctant to label their experiences as sexual harassment, or perceive the harassment behaviors to be normal or expected.

4. INCIDENCE OF SEXUAL HARASSMENT

According to generally accepted statistics, approximately 50% of all women in the United States will experience sexual harassment at some time either in the workplace or in educational settings. Because initial American studies focused primarily on White women, those data are presented first, followed by a consideration of sexual harassment of men, minority women, and cross-cultural victims.

4.1. U.S. Incidence Data on Sexual Harassment of White Women

4.1.1. Workplace Incidence Data

Although certainly not all sexually harassed women have experienced the level of harassment reported by Michele Vinson, who was forced to have sexual relations with her boss 40 to 50 times during her employment at the Meritor Savings Bank (*Meritor v. Vinson* 1986 court case), the incidence of sexual harassment for American White women is high. Three often-cited landmark studies performed by the U.S. Merit Systems Protection Board (USMSPB) yielded similar incidence rates of sexual harassment of White women. A stratified random sample of 23,964 federal workers in the 1981 USMSPB survey revealed that approximately 42% of the female respondents had experienced sexually harassing behaviors over the previous 2 years. The most common form of harassment consisted of verbal comments or sexual remarks (33%), followed by physical touching (26%), and being pressured for dates (15%), with nearly 10% having been pressed for sexual cooperation and a small percentage having experienced sexual assault. Many women reported having been victimized repeatedly over a long period of time. Young and single women were the most likely targets of harassment behavior.

In another classic, frequently cited study, in 1985, Gutek found that 53% of the women in a community sample that was representative of the workforce had been sexually harassed at least once. Of these respondents, 19.8% had experienced insulting comments, 15.0% had experienced insulting looks and gestures, 24.2% had experienced sexual touching, and 7.6% had been expected to engage in sexual activity as part of their employment. Women in male-dominated occupations experienced sexual harassment at a higher rate than that of women in other occupations.

At least two major studies have been conducted by the U.S. military. In 1991, the U.S. Navy administered the Naval Equal Opportunity Sexual Harassment (NEOSH) survey, which revealed that 44% of enlisted women and 33% of women officers had been sexually harassed, figures comparable to those in the USMSPB studies. More recently, the Department of Defense surveyed 90,000 personnel in all branches of the armed forces. In this survey, 78% of the women had at least one sexual harassment experience. Women in the marine corps had the highest incidence rate in this study (87%), whereas those in the air force had the lowest incidence rate (74%).

4.1.2. Academic Sexual Harassment

It is difficult to determine the actual incidence rate of academic sexual harassment because the figures from various studies are quite discrepant, ranging from 13 to 70%. Explanations for these diverse statistics include differences in the type of survey (e.g., specific behavioral questions, a general question about whether or not the respondent has been sexually harassed), varying response rates, representativeness of the educational institution (e.g., most surveys are completed at 4-year colleges or universities and do not sample community colleges), and whether the focus is on faculty harassment or peer harassment. However, even conservative estimates suggest that from 30 to 50% of women who have completed these surveys have been harassed, a figure that corresponds to workplace statistics. Once again, the frequency of milder forms of sexual harassment is higher than the occurrence of sexual coercion. For example, in one study, approximately 30% of the women were subjected to gender harassment, 15% to unwanted sexual attention, and 9% to sexual coercion.

Evidence suggests that sexually harassing behavior may begin at the high school level. A classic study by the American Association of University Women (AAUW), questioning students from the 8th through 11th grades, reported that 85% of the girls (87% of the White girls) had experienced sexual harassment, with 80% of the harassment emanating from male peers and 20% from male adults associated with the school. This figure may be somewhat inflated because methodological issues have been raised about the procedure of the survey, but other research supports the contention that high school sexual harassment is a significant problem that is just beginning to be addressed in the literature. It is important to determine the incidence of peer harassment in this setting because patterns of sexually harassing behaviors that are developed and reinforced as a form of "male bonding" in the high school setting may continue at the university level and in the workplace.

4.2. Specific Victim Groups

4.2.1. Men as Victims of Sexual Harassment

In the USMSPB studies, from 14 to 19% of men reported being sexually harassed in the federal workplace. In the Department of Defense study mentioned earlier, 38% of men reported sexual harassment experiences, with 35% having experienced gender harassment and less than 1% having experienced attempted

rape. In addition, the AAUW high school study included the statistics that, overall, 76% of boys had been sexually harassed, with 56% having been the recipients of sexual comments and 42% having been touched in a sexual manner.

The impetus for the recent interest in men as sexual harassment victims stems from the 1998 Supreme Court decision in *Oncale v. Sundowner Offshore Services* that stated that same-sex harassment meets the requirement for sexual harassment cases in that the victim is singled out due to his gender. This behavior was classified as producing a hostile work environment regardless of whether the plaintiff in a civil suit was the target of sexual assault (as was the case in the Joseph Oncale lawsuit), was intimidated by male coworkers, or had his masculinity questioned in severe and persistent incidents. Related evidence in the USMSPB study indicated that approximately 22% of the male harassment victims identified in the survey had been sexually harassed by other men (37% in the 1988 study), thereby constituting 91% of all same-sex harassment victims. In the Department of Defense study, 51% of the male respondents reported same-sex harassment experiences as compared with 2% of the women, although overall, women had experienced more than twice as much harassment (76%) as had men (36%). Another study showed that 44% of the men were targets of same-sex harassment, with the majority of the harassment incidents involving "crude/offensive behaviors" such as jokes, gestures, and derogatory remarks about men. Recent studies have used the Sexual Harassment of Men (SHOM) scale, a revised version of the SEQ-R, to identify men who have been sexually harassed. Although male victims generally seem to have fewer negative reactions to their experiences, a sizable percentage of men have experienced hostile environment conditions in the workplace, with the primary form of same-sex harassment being related to attacks on the masculinity of the male targets. The actual incidence rate of male victims may be higher than has been indicated so far given that men might be reluctant to report being victimized due to fear of not being believed or possibly because complaining might go against stereotyped views of appropriate male behavior.

4.2.2. Women of Color

Few studies have been conducted to examine sexual harassment incidence figures of women of color as compared with those of White women, and the pattern of findings in these studies has been inconsistent. Some studies concluded that women of color have greater experiences of sexual harassment than do White women, and others illustrated the opposite pattern. Earlier studies often did not provide incidence data broken down by race or ethnicity and often omitted economic status information. If a woman of any race has a low income and low status, these factors may increase her risk for becoming a sexual harassment victim. In addition, it is important to understand that for women of color, sexual harassment (as a form of sex discrimination) and race discrimination often are intertwined. The issue becomes even more complex when women of color are sexually harassed by men of color. For example, women of color had conflicting reactions to the testimony of Hill against Thomas in his Senate confirmation hearings. Although they might identify with Hill as a woman of color, it is possible that their racial identification was more important than gender identification in this situation, and they also may have been reluctant to see Thomas, a Black man, being prevented from succeeding due to Hill's testimony. Another important factor that must be considered is that the category of women of color is too diverse. For example, African American, Hispanic American, and Asian American women may experience differential rates of sexual harassment and respond in various ways to this harassment.

4.2.3. Cross-Cultural Victims

The methodological issues that exist for cross-cultural comparisons of sexual harassment incidence rates are similar to those described in the preceding section on women of color, but on a larger scale. Lack of information on incidence rates in the countries sampled, use of convenience samples that are not representative, lack of a definition or even the concept of sexual harassment in some countries, use of a variety of measures, and translation issues all are barriers to a definitive comparison of incidence rates cross-culturally. However, these comparisons are significant not only because harassment is a worldwide problem but also due to the presence of multinational companies in various countries. For example, a charge of sexual harassment was directed against a male Japanese plant manager at an American subsidiary of a Japanese company. The American woman employee, who was a supervisor, alleged that the manager had created a hostile work environment by inviting her to dinner several times, giving her gifts, massaging her knee after a ski injury, and trying to kiss her. The manager denied the allegations and explained

his behavior by stating that he always invited groups of workers to dinner, that giving gifts in appreciation was a Japanese custom, that he never tried to kiss her, and that he only wanted to stop her pain by massaging her knee. The company settled her claim, although the president suggested that the manager was just trying to follow orders to become friendly with American workers and that no sexual harassment was involved. This case illustrates the confusion that can arise in identifying sexual harassment in multicultural settings.

This subsection includes data from a sampling of countries. Canada generally has demonstrated incidence rates similar to those of the United States, as have countries in Northern Europe. An International Labor Office survey determined that 30% of Austrians sampled, 13 to 58% of Dutch respondents, and 21% of French citizens had experienced sexual harassment. Although figures vary from 30 to 80% in Germany, a conservative estimate would be that 59% of German women have experienced sexual harassment. In Belgium, 34% of women and 12% of men indicated that they had been sexually harassed, whereas the incidence rate was 48% for Italian respondents and 36% for Finnish respondents, with police respondents reporting the highest rate. Victimization in the United Kingdom ranged from 47 to 90%, with women in the police once again subjected to the greatest amount of harassment.

Data from Far Eastern countries exhibit somewhat higher patterns of harassment. A 2001 survey in Japan indicated that approximately 75% of all respondents had experienced sexual harassment at least once, and 70% of public officers in Korea reported having been sexual harassment victims. However, in 1995, a somewhat low rate of 25% of Hong Kong university students reported that they had been subjected to sexual harassment. In one of the few studies conducted in Pakistan, incidence rates revealed that 58% of nurses, 95% of agricultural workers, 91% of domestic workers, and 93% of women employees in private and public organizations had been victims of sexually harassing behaviors. In India, researchers suggested that harassment was widespread and was actually encouraged due to societal beliefs that men are naturally sexual "predators," women are sexual "prey," and so harassment was expected and considered acceptable.

Little data could be obtained from the Middle East (e.g., Israel was seen as having somewhat similar but possibly higher incidence rates compared with those in the United States, Turkey was described as a society in which harassment might be extensive due to its patriarchal nature), Latin America (e.g., Brazil was portrayed as having similar rates of gender harassment and sexual coercion, but higher rates of unwanted sexual attention, compared with those in the United States), or Africa (e.g., there are no government statistics on incidence rates in South Africa).

4.3. Characteristics of Male Harassers

The major psychometric measure in this area is the Likelihood to Sexually Harass (LSH) scale, developed by Pryor in 1987, that describes 10 scenarios where there is a potential for engaging in sexual harassment without punishment. Men who score high on the LSH scale are portrayed as having traditional masculine sex roles. The classic type of harasser is the boss who engages in "sexual exploitation" or uses his power to gain sexual favors. Another type is a socially inept man who misinterprets women's friendly cues as indications of sexual interest. A third type that is more prevalent in male-dominated workplaces identifies men who are hostile toward women and who see the workplace as the in-group (us or men) against the out-group (them or women). Finally, men who are opposed to homosexuality (i.e., homophobes) may engage in harassment of women and men. Researchers have proposed that men who score high on the LSH scale are more likely to engage in harassment behaviors in workplace settings where harassment is accepted as normative behavior.

5. OUTCOMES AND RESPONSES

It is extremely important to examine the consequences of sexual harassment victimization as well as the coping responses that victims use to improve educational efforts in prevention programs, to inform expert testimony in sexual harassment cases, and to train therapists who treat victims of sexual harassment.

5.1. Consequences of Sexual Harassment

One analysis suggested that victims may experience the "sexual harassment trauma syndrome," which consists of the following components.

5.1.1. Emotional Reactions

Victims often are angry, embarrassed, and fearful, and they experience learned helplessness or feelings of powerlessness.

5.1.2. Physical Outcomes

These outcomes may include sleeping disruptions, headaches, phobias, and fatigue.

5.1.3. Alteration in Self-Perception

Victims often suffer from lowered self-esteem and self-confidence. For example, employees may perceive themselves as being promoted due to their sexual attractiveness rather than their competence.

5.1.4. Behavior Change

If victims feel responsible for the harassment, they may try to dress more conservatively and wear "less sexy" clothes. Their trust in men often is adversely affected, and they may avoid interpersonal relationships. As a result, friendship and alliances with coworkers might not develop, and victims often are excluded from mentoring relationships that can enhance their careers. In terms of academic sexual harassment, students might decide to change careers, or at least have their research experiences curtailed, as a result of their harassment experiences.

The integrative theory described earlier proposes that sexual harassment victims experience "work-related withdrawal" or attempts to remove themselves from the job environment. These effects can include lowered productivity, decreased concentration on work-related tasks, absenteeism, and being late for work. Another consequence may be "job-related withdrawal," which refers to individuals' plans to retire early or quit their jobs. In addition, the theory predicts that health-related effects will occur, similar to the physical outcomes described previously, as will damaging psychological effects from depression and anxiety. Although sexual coercion clearly will produce more extensive negative effects than will milder forms of harassment, there is evidence that the cumulative effects of ongoing persistent sexual harassment of any type will have a deleterious impact on the victim.

This "poisonous climate" of sexual harassment may affect the organization or educational institution by lowering organizational commitment and damaging the reputation of the company or educational institution. The cost of sexual harassment to the organization or educational institution can be severe and includes turnover, increased medical costs, low morale, lowered productivity, and absenteeism. It generally is impossible to determine the actual human and monetary costs of this behavior, although the USMSPB's 1988 study suggested that sexual harassment may have cost the U.S. government $250 million over 2 years. Of course, if sexual harassment complaints reach a lawsuit level, the costs can be enormous.

5.2. Coping Responses

One of the common reactions to the testimony of Hill, alleging that she had been sexually harassed by Thomas several years earlier, included "Why didn't she just leave?" and "Why didn't she tell him to stop the behavior?" These responses represent direct coping attempts to eliminate the harassment. If a victim does not behave according to these public assumptions, the situation might be judged as insufficiently offensive or might not be acknowledged as a case of sexual harassment. The frequency of use of different types of coping and the effectiveness of these various coping responses must be examined to assist harassment experts in advising victims. In addition, if researchers clearly indicate that most victims do not report harassment incidents and that victims often endure these incidents for long periods of time without any direct response (as was the case with Hill), organizations may be persuaded to adopt procedures that will discourage harassment and encourage victims to report any harassment to supervisory personnel.

5.2.1. Types of Coping Responses

Although several category systems exist, the most useful approach divides the type of response into direct (active) versus indirect (passive) reactions to sexual harassment incidents. The most basic direct reaction to such an incident is to tell the harasser to stop the behavior in a face-to-face confrontation or in a letter. The next step is to report the behavior to a supervisor. The victim then may escalate the process to filing a formal complaint, although on the average only 1 to 7% of victims reach this stage. If these processes do not stop the harassing behavior, the final direct approach is to institute a lawsuit.

A number of indirect or passive responses to harassment have been noted. A victim may ignore the

behavior, trivialize the incident, redefine the behavior (e.g., "He's just fooling and doesn't mean it"), and/or attribute it to the harasser's unhappy marriage or the victim's provocative clothing. Alternative passive responses include avoiding the harasser, redirecting the conversation away from personal comments, and bringing a friend or colleague to meetings with the harasser. Avoiding the harasser is difficult if the perpetrator is a supervisor or professor. A victim may also seek social support by venting to friends and family. Finally, if the behavior continues, the individual may quit her job, transfer to another department, or drop the course (in the case of academic sexual harassment).

The most common response of victims is to ignore the behavior or attempt to avoid the harasser. However, harassers generally continue to behave inappropriately, and if no action is taken against these individuals, the harassment may escalate. Sexual harassment experts usually advise victims to respond directly to harassers, but there is a disconnect between this advice and the behavior of sexual harassment victims. The major reason why victims do not respond assertively is fear of retaliation or the perception that this is a high-risk coping response. Victims often are in low power positions and are afraid of losing their jobs if they complain about harassment, especially if the harassers are in supervisory roles. For example, a woman in a relatively high power position told her harasser to "cut it out," but when asked why she did not report him, she said that she wanted to hold onto her job. Studies indicate that using direct confronting types of coping styles often backfires for the employees in that one-fourth to two-fifths of individuals who complained were forced from their jobs. In one survey, 75% of the victims who directly confronted their harassers found that the behavior continued after these confrontations. In fact, some evidence suggests a paradox for victims. If they respond more assertively, there is a possibility that the behavior will stop and the harasser will be punished, but these victims also may experience more negative job-related outcomes (e.g., in terms of others' impressions of them and their commitment to the job) than will those who respond more passively.

A second related reason why assertive coping often does not occur is that complainants often are labeled as "troublemakers" or treated as "whistleblowers." For example, a female junior faculty member who reported harassment incidents by a male senior professor was told that he was a more valuable employee and that she could leave if she intended to persist in her complaint. In Japan, where "saving face" is a core value, victims often are blamed for their harassment experiences and are seen as bringing shame to their companies' reputations and preventing the harassers from succeeding in their positions. When an employee goes forward with a complaint, she is suspended without pay while the accused aggressor remains in his job.

Another interpretation suggests that if the climate is supportive of aggressors, victims may avoid reporting this type of behavior because they believe that nothing will be done to stop the harassers. Finally, there are individual differences in the tendency to use direct coping styles, with some victims feeling too embarrassed or humiliated by their experiences to lodge complaints. Sexual harassment experts should reexamine their advice to victims considering the tendency of employees to adopt nonassertive responses to the offensive behavior.

6. PREVENTION AND INTERVENTION

As a result of increased public attention to the issue, most major companies have developed policies against sexual harassment and a number of consulting firms are specializing in training programs to prevent sexual harassment. Prevention is the preferred approach to eliminate harassment in the workplace and educational settings because once inappropriate behavior occurs, it results in a "lose–lose" situation for all concerned. No matter how the situation is resolved, the company will incur human and financial costs as well as potential damage to its reputation, the victim will experience negative outcomes, and the accused harasser, if found to be guilty of sexual harassment, will be subject to sanctions that may extend to termination of employment. In addition, the company could even be sued by the victim for not preventing the harassment and by the defendant for defamation of character or for terminating him without "due process."

If top management is only motivated to protect itself from lawsuits, a simple broad policy statement may be the extent of its prevention attempts. However, if top management truly wants to create a harassment-free culture or environment, the approach will be more extensive and will involve employees at every level. Even with the best of intentions, there are no sure-fire programs that can guarantee a positive climate in the workplace. In fact, research suggests that "one size does not fit all"; in other words, programs must be developed

within the framework of specific companies. Another crucial element is evaluation of the effectiveness of the prevention efforts. To date, critics have complained that most training programs have not been established as valid in terms of prevention. One of the difficulties is to determine the outcome measure. If the program is effective, will the number of complaints increase given that victims will feel more comfortable in filing complaints, or will the number decrease because sexual harassment behaviors have been prevented?

One approach suggests that there are three important stages of prevention and intervention in sexual harassment situations. The first relates to policies explicitly stating that sexual harassment will not be tolerated in the organization. The policy statement also should include specific examples of inappropriate behavior with associated penalties, should provide contact sources for victims, and should perform an educational function. Most important, this policy should be disseminated to all employees (e.g., through handouts to new employees when hired as well as through systematic distribution of the guidelines to other employees) and supported from top management down throughout the hierarchy. Unfortunately, policy statements alone are not effective, particularly if the climate or culture supportive of sexual harassment is not altered.

The second component is procedures that have been developed as part of the intervention process. Employees must be informed that if a formal complaint of sexual harassment is brought against an individual, both the rights of the accused and the accuser will be protected. At this point, a "neutral" investigator (an in-house employee or an outside consultant) should be in place to determine the validity of the charges.

As discussed earlier, however, the formal complaint route probably will not be chosen by many victims due to fear of retaliation or fear of being labeled as a "troublemaker." In addition, victims might not want their harassers to be sanctioned or punished. Therefore, the informal complaint approach might be more effective in some settings. A neutral negotiator is identified for victims, and face-to-face meetings may be arranged between victims and perpetrators. In many cases, all the victims want is for the harassment to stop. Possible outcomes of this informal process include stopping of the behavior with or without an apology, transfer of one or both of the individuals, and (in extreme cases) resignation of one or both of the individuals. Informal procedures have been criticized for lack of standardized objective procedures that may lead to ineffective resolution of the problem situation.

The third step in the intervention process is training given that policies and procedures alone often have been shown to be ineffective in preventing harassment. Training should be used to sensitize employees to types of sexual harassment and effects on victims, should include sanctions for harassment behaviors, and should occur over extended periods of time. These programs should include not only an informational component but also interactive exercises such as role-playing and watching videos with actors, followed by intensive discussions. Training programs have been shown to increase understanding of the behaviors constituting harassment but might not prevent these behaviors. In addition, if programs are voluntary, the people attending might not be the potential harassers, and if programs are mandatory, resentment and resistance might occur. In one study, for example, males scoring high on the LSH scale showed a small boomerang effect after watching a training video.

Creative and idiosyncratic approaches may effectively supplement the three basic stages described previously. At one university, in addition to training films, lectures, and a "sexual harassment awareness week," core faculty and staff volunteered to be resource individuals for students who had experienced sexual harassment. This group was trained not only in referral sources and policies and procedures but also in communication skills so that the faculty could counsel students and be supportive during these stressful situations. Once again, it is important for organizations to determine appropriate ways in which to implement these three steps and to develop effective means of preventing harassment.

7. CONCLUSION

Sexual harassment is a worldwide problem that causes distress to its many victims and damages the various organizational climates. More attention should be directed at developing adequate prevention and intervention approaches that contain an evaluation component to establish the effectiveness of these programs.

See Also the Following Articles

Coping ■ Feminist Psychology ■ Gender and Culture ■ Gender Role Development ■ Homophobia ■ Homosexuality ■ Occupation and Gender ■ Prejudice and Discrimination ■ Risk-Taking and Sexual Behavior ■ Sexual Behavior ■ Sexual Behavior and Culture

Further Reading

Blumenthal, J. A. (1998). The reasonable woman standard: A meta-analytic review of gender differences in perceptions of sexual harassment. *Law and Human Behavior*, *22*(1), 33–57.

Equal Employment Opportunity Commission. (1980, April). *Guidelines on discrimination because of sex under Title VII of the Civil Rights Act as amended: Adoption of interim interpretive guideline*. Washington, DC: Government Printing Office.

Fitzgerald, L. F. (1993). Sexual harassment: Violence against women in the workplace. *American Psychologist*, *48*, 1070–1076.

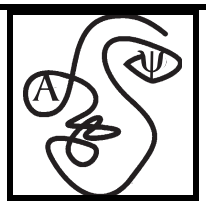
Fitzgerald, L. F., Drasgow, F., Hulin, C. L., Gelfand, M. J., & Magley, V. J. (1997). Antecedents and consequences of sexual harassment in organizations: A test of an integrated model. *Journal of Applied Psychology*, *82*, 578–589.

Fitzgerald, L. F., Gelfand, M., & Drasgow, F. (1995). Measuring sexual harassment: Theoretical and psychometric advances. *Basic and Applied Social Psychology*, *17*, 425–445.

Gutek, B. A., & Koss, M. P. (1993). Changed women and changed organizations: Consequences of and coping with sexual harassment. *Journal of Vocational Behavior*, *42*, 28–48.

O'Donohue, W. (Ed.). (1997). *Sexual harassment: Theory, research, and treatment*. Needham Heights, MA: Allyn & Bacon.

U.S. Merit Systems Protection Board. (1995). *Sexual harassment in the federal workplace: Trends, progress, and continuing challenges*. Washington, DC: Government Printing Office.



Social Comparison and Subjective Well-Being

Thomas Ashby Wills and Don Mendoza

Albert Einstein College of Medicine, Bronx, New York, USA

1. Is Social Loafing a Universal Effect with Disastrous Consequences for Group Performance?
2. A Brief Historical Overview
3. Theoretical Explanations
4. Implications of Social Loafing
5. Prevention and Control of Social Loafing

GLOSSARY

downward comparison Comparing one's own attributes with those of another person who is worse off on the attributes.
self-enhancement Using social information to feel better about one's own situation.
self-evaluation Using social information to obtain an accurate assessment of one's own standing.
self-improvement Using social information to improve one's own performance.
threat-derogation effect Making more negative ratings of other people's attributes when one is threatened, in order to increase perceived social distance between the self and the other.
upward comparison Comparing one's own attributes with those of another person who is better off on the attributes.

Social comparison research considers how people's perceptions of their own situation are influenced through comparison with the attributes and situations of other people. Research has identified processes of upward comparison and downward comparison, and studies have linked these to subjective well-being. Downward comparisons may serve to increase subjective well-being, particularly in the short term after people have experienced setbacks or failures. Upward comparison can be used for self-improvement, but evidence from field studies suggests that a constant background of unfavorable comparisons is negatively related to psychological and physical outcomes. Costs and benefits of comparison processes, and their implications for intergroup relations, are discussed in this article.

1. SOCIAL COMPARISON THEORY

Subjective well-being (i.e., feelings of positive mood and life satisfaction) is an important variable for applied psychology, both for its theoretical interest and for its relations with health and productivity. Social comparison theory is relevant for the study of subjective well-being because comparisons with other people can influence how people feel about themselves and their own situation. Demographic characteristics such as gender and ethnicity are not strongly correlated with well-being indices, so research has shifted toward considering basic psychological processes that

©2004 Elsevier Inc.
All rights reserved.

contribute to lower vs higher levels of subjective well-being. Social comparison is one of these processes.

Social comparison theory was developed to address the question of how people assess their own attributes and abilities when objective measures are not available. For some characteristics (e.g., weight) objective measures are readily available, but this is not so much the case for important personal attributes (e.g., attractiveness). For such attributes, when people ask themselves "Where do I stand on attribute X?" an answer must come from comparing the self with others. The form of the answer, then, is "Comparing myself with the attribute for persons A, B, C, and D, my standing is most like that of B." The outcomes of these comparisons can have important implications for whether people feel happy or unhappy about themselves and their own situation.

1.1. Upward Comparison

The original formulation of social comparison theory by Leon Festinger was based on the assumption that people want an objective evaluation of their own attributes (e.g., academic competence). This led to the concept that people would tend to compare themselves with others who were better off than the self on the attribute in question, termed upward comparison. The predicted consequence of such comparisons was that uncertainty about one's own standing would be reduced and therefore one would feel more certain about performance (or self-improvement) in this domain. Support for these predictions was obtained by studies that showed a preference for comparison with slightly better-off others, mostly from contexts in which the subjects were already doing pretty well on the attribute in question and felt confident about attaining the status of the better-off comparison others. Further developments by Goethals and Darley led to the recognition that in the comparison process, people would also consider additional attributes that are believed to be related to the attribute in question (e.g., intelligence and time spent studying).

1.2. Downward Comparison

The issue for this model of human functioning was what would happen when people believed that they were not doing so well on the attribute in question, a context that could be true for part of the population. This led to the formulation by Wills of downward comparison theory, which considers the types of comparisons that occur in situations in which people are not so well off to begin with. For this context, the prediction was that people would tend to compare with others who are worse off than the self, termed downward comparison. The predicted consequence of this type of comparison was that people would feel better about themselves and their own situation. Support for this theory was obtained in situations in which people felt that their own status was poor or threatened. For example, downward comparisons were found to be common among people with physical illness, who reported ways in which their own situation was relatively favorable compared to others, and media studies showed that people felt better about their own city when the local newspaper portrayed a high level of crime in other cities. Further development by social psychologists such as Taylor and Lobel led to recognition that some downward comparisons may be with cognitively constructed images of typical others, and cognitive comparisons may have different consequences than forced contact with other people having problems.

2. RESEARCH ON PRINCIPLES OF SOCIAL COMPARISON

A number of laboratory studies have tested propositions from social comparison theory. Because social comparison is a cognitive process and hence not directly observable by a third party, findings from laboratory research have been helpful for understanding comparison processes because known variables can be experimentally manipulated and conclusions about causality can be drawn from the observed effects of these variables.

Results from laboratory studies typically show a downward shift in comparison preferences when distress, threat, or failure is present. When people feel they are doing well, they tend to prefer information about others who are better off than the self. However, when negative mood is experimentally induced or when subjects are identified who have experienced a recent setback (e.g., in academic performance or social relationships), then there is a preference for information about others who are having similar problems or worse problems. This is consistent with the basic principles of downward comparison theory. Another aspect of social comparison is that threatened people may work to improve their own self-image through

derogating the characteristics of other people or groups, a process that represents the potential dark side of social comparison and has been termed the threat-derogation effect. This process has been demonstrated consistently among threatened people in laboratory settings, and research on intergroup perceptions by Brewer and others has established this effect in downward comparisons with outgroups.

The predicted consequences of social comparison have also been observed in the laboratory. When distressed people get access to information about worse-off others, they typically show an improvement in subjective well-being—that is, reduced negative affect or increased positive affect. Effects of upward comparison tend to depend on the situation, but several studies have shown that when subjects are exposed to information about others who are better off (more attractive, more intelligent, better adjusted, etc.), the subjects show increases in negative affect. The primary moderator of upward comparison effects is the controllability of the problem. When people feel that a difficulty is easily solvable and they can engage in self-improvement, they tend to react more favorably to upward comparisons; however, when they perceive their problems are not controllable, they react more negatively.

Laboratory research has also clarified that there are several possible motives for social comparison. Most common are self-enhancement, using comparisons to obtain a favorable picture of the self (e.g., I compare "to feel good about my own situation"); self-assessment, using social information to obtain a strictly accurate evaluation of one's relative standing in a social group (e.g., "to see how I'm doing"); and self-improvement (e.g., "so I can get better"). These motives are all present in the population to varying degrees and may be activated under different conditions. This research has been extended in applied settings (e.g., with business entrepreneurs) to consider how social comparison motives for making money are related to subjective well-being.

There is also evidence of consistent individual differences in social comparison. For example, some studies indicate that women engage more often in all types of social comparison. Other research has indicated dispositional aspects of social comparison. A dimension called comparison interest indicates that some people are more interested, in general, in pursuing social comparisons. Also, a dimension termed dispositional envy has been identified, in which some people are particularly disposed to feel envious toward better-off others, and hence unhappy, because of comparison processes.

3. RESEARCH ON SELECTED AND FORCED COMPARISONS IN FIELD SETTINGS

3.1. Selected Comparisons

Research in field settings has examined several questions about how social comparisons are related to subjective well-being. A number of studies have been conducted on people subjected to distress from physical illness, such as heart disease or cancer. Downward comparisons are frequent in these populations, with people reporting that they can think of others who are worse off in some respect. When patients are studied in clinic settings, in which a range of comparison others are available, a paradoxical aspect is also noted to social comparison in these populations. Helgeson and others have shown that patients show a preference for affiliating with other patients who are recovering from the disease, apparently because this makes them more optimistic about their own recovery. At the same time, however, they tend to compare cognitively by thinking about others who are worse off, and this is related to greater subjective well-being as indicated by feelings of current mood and perceived coping ability.

Studies of individuals with mental illness are fewer in number and still paint a somewhat confusing picture. People with clinical depression, whose basic issue is that they feel very negative about the self, seem less inclined to engage in active, self-aggrandizing comparisons, possibly because of a breakdown in self-enhancement processes. However, there is evidence that even moderately unhappy people find ways to engage in indirect forms of self-enhancement, and can be creative about identifying unique aspects of the self on which they perceive themselves as advantaged relative to others, as a way of maintaining some level of self-regard. It has been suggested that depression may be brought on by excessive reliance on social information and an inability to discount (or ignore) unfavorable comparisons, but there has been no longitudinal research in this area and the answer to the question remains unresolved.

3.2. Forced Comparisons

The other question about social comparison in real-world settings concerns what happens when people are repeatedly exposed to comparison information that they cannot avoid. One example of forced comparison is school settings, in which students receive a

considerable amount of information about others' performance, either through public postings and awards or through more informal communication. Findings from studies by Marsh and others have been consistent in showing that students feel worse about themselves, in terms of academic self-concept, when they are in schools in which many students are performing better than they are performing. Conversely, people feel better about themselves when they are in school environments in which most students have only average performance. The suggestion has been that constant exposure to unfavorable comparisons makes people feel less efficacious and less inclined to exert effort for self-improvement. These findings have been summarized as indicating the "big fish, little pond" effect—the concept that in terms of subjective well-being it can be better to be a relatively big fish in a small pond, even if one does not learn to swim quite as well.

Another noteworthy question is how people respond to comparisons of income, an attribute that may not be publicly posted but is still clearly evident in daily life. Studies using evidence from aggregate statistics and public health data have been consistent in showing income distributions in the population related to individuals' reports of subjective well-being. For example, reports of happiness and life satisfaction are decreased when there is greater inequality in distributions of income (i.e., when the average individual sees more people with higher income than him or her). This has been demonstrated both within the United States and across different countries. Also, evidence from studies using several different indices of inequality has shown that mortality rates are higher in census units where there is greater inequality of income. Thus, forced comparisons have definite consequences for subjective well-being.

4. CONCLUSIONS AND IMPLICATIONS

Social comparison is an important contributor to subjective well-being. Laboratory studies have established basic social comparison principles in controlled settings, and field studies have tested how these processes operate in the natural environment. This research has shown that when people perceive they are doing well they tend to engage in upward comparison, but when people are stressed or threatened there is a downward shift in comparison processes. Downward comparison has been supported as a selective process of using comparison information so as to enhance subjective well-being. It is probably a mixed blessing as a coping strategy, if exclusive reliance on downward targets serves in the long term to deter self-improvement. The available theory and evidence suggest that downward comparison is best as a coping strategy when used judiciously in times of distress so as to maintain self-esteem and provide a bridge to sustained, active coping efforts.

Comparison with better-off others has been supported as a method to self-improvement, but aspects of the research suggest that upward comparison can be a double-edged sword. Findings from field studies have shown forced upward comparisons to be related to lower levels of psychological and physical well-being. Laboratory research suggests that the comparison process may become so automatized that people may not always be aware of the impact of social comparisons, particularly when they are cognitively busy dealing with the demands of daily life. This work suggests that people can sometimes "undo" a social comparison by discounting the relevance of the comparison dimension, but a constant background of unfavorable comparisons could be difficult to undo and could work to decrease subjective well-being. Linkages of forced comparisons to physical health have been identified, but the processes underlying these effects are not well understood and implications of this evidence for applied psychology need to be considered.

Social comparison research also has applied implications because of linkages to intergroup relations. The concern is that the threat-derogation effect may be operative among people whose subjective well-being is threatened because they believe that they are not doing well, their problems are uncontrollable, and their situation may not improve. This concern is bolstered by archival studies that show stereotyping and authoritarian attitudes in the U.S. population to be increased in times of threat (e.g., from economic recession). Whether this effect is related to derogatory attitudes toward endogenous targets (e.g., ethnic minorities), or other groups such as immigrants, has not been extensively studied in applied research in field settings and remains a topic of importance for investigation.

See Also the Following Articles

Well-Being

Further Reading

Banaji, M. R., & Prentice, D. A. (1994). The self in social contexts. *Annual Review of Psychology*, *45*, 297–332.

Bogart, L. M., & Helgeson, V. S. (2000). Social comparisons among women with breast cancer. *Journal of Applied Social Psychology*, *30*, 547–575.

Brewer, M. B., & Weber, J. G. (1994). Self-evaluation effects of interpersonal versus intergroup social comparison. *Journal of Personality and Social Psychology*, *66*, 268–275.

Buunk, B., & Gibbons, F. X. (Eds.). (1997). *Health, coping, and social comparison*. Mahwah, NJ: Erlbaum.

Hagerty, M. R. (2000). Social comparisons of income in one's community: Evidence from national surveys of income and happiness. *Journal of Personality and Social Psychology*, *78*, 764–771.

Kawachi, I., Kennedy, B. P., Lochner, K., & Prothrow-Stith, D. (1997). Social capital, income inequality, and mortality. *American Journal of Public Health*, *87*, 1491–1498.

Lyubomirsky, S., & Ross, L. (1997). Hedonic consequences of social comparison: A contrast of happy and unhappy people. *Journal of Personality and Social Psychology*, *73*, 1141–1157.

Marsh, H. W., Kong, C.-K., & Hau, K.-T. (2000). Models of the big fish–little pond effect on academic self-concept. *Journal of Personality and Social Psychology*, *78*, 337–349.

Schwarz, N., & Strack, F. (1999). Reports of subjective well-being: Judgmental processes and their implications. In D. Kahneman, E. Diener, & N. Schwarz (Eds.), *Well-being: The foundations of hedonic psychology* (pp. 61–84). New York: Russell Sage Foundation.

Srivastava, A., Locke, E. A., & Bartol, K. M. (2001). Money and subjective well-being: It's not the money, it's the comparison motives. *Journal of Personality and Social Psychology*, *80*, 959–971.

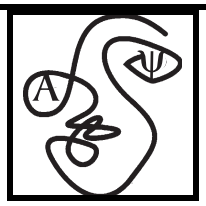
Suls, J., & Wheeler, L. (Eds.). (2000). *Handbook of social comparison*. New York: Kluwer Academic/Plenum.

Taylor, S., & Lobel, M. (1989). Social comparison activity under threat: Downward evaluation and upward contacts. *Psychological Review*, *96*, 569–575.

Wills, T. A. (1981). Downward comparison principles in social psychology. *Psychological Bulletin*, *90*, 245–271.

Wills, T. A., & Sandy, J. M. (2001). Comparing favorably: A cognitive approach to coping through comparison with other persons. In C. R. Snyder (Ed.), *Coping with stress: Effective people and processes* (pp. 154–177). New York: Oxford University Press.

Wood, J. V., & Lockwood, P. (1997). Social comparisons in dysphoria and low self-esteem. In R. Kowalski & M. Leary (Eds.), *The social psychology of emotional and behavioral problems* (pp. 97–135). Washington, DC: American Psychological Association.



Social Loafing

Francisco Gil

Complutense University, Madrid, Spain

1. Is Social Loafing a Universal Effect with Disastrous Consequences for Group Performance?
2. A Brief Historical Overview
3. Theoretical Explanations
4. Implications of Social Loafing
5. Prevention and Control of Social Loafing
Further Reading

GLOSSARY

additive tasks Tasks in which members' inputs are added to produce a group result (e.g., generating the highest number of ideas in a brainstorming session); group productivity depends on all members' contributions.

coactive tasks Tasks in which people work individually in the presence of others who are performing the same task (e.g., doing administrative work in a collective office).

conjunctive tasks Tasks that must be performed by all group members (e.g., team cycling competitions); group productivity depends on the worst group member (e.g., the slowest cyclist).

disjunctive tasks Tasks that may be performed by one member of the group (e.g., solving a problem); group productivity depends on the best (in this case the most expert) group member.

free riding When individuals who have not made appropriate contributions to the task profit from the contributions of other group members.

social facilitation An increase of individual effort when individuals perform tasks in the presence of other people or alongside coactors; this effect tends to occur when highly familiar or overlearned tasks are performed and the correct responses are dominant.

social loafing A reduction of individual effort when individuals perform collective tasks, as compared with working individually or alongside coactors.

sucker effect When individuals decide to reduce their own effort and input to group performance (and hence diminish the reinforcements they could receive) instead of being "suckers" and letting themselves be exploited by free riders.

The social loafing effect is related to the loss of motivation that occurs in a group context, causing people to reduce their effort or to produce less when they carry out collective tasks than when they work alone. This effect is the opposite of the social facilitation effect, which is related to the increased achievement that occurs when performing individual tasks in the presence of others. Social loafing is related to two other effects that occur in certain enhancing conditions: the free rider effect, which occurs when individuals who have not made an appropriate contribution to the task profit from the inputs of other group members, and the sucker effect, which occurs when individuals decide to reduce their own effort to group performance rather than let themselves be exploited by free riders. Different theories (e.g., social impact, arousal reduction, identification and evaluation potential, dispensability of effort, self-efficacy, collective effort) have questioned the original supposition that social loafing is a universal process inherent to work in groups and have identified

©2004 Elsevier Inc.
All rights reserved.

different conditions in which it is more likely. Various factors, mainly those related to tasks, groups, and organizations, have also been found to prevent and control the social loafing effect.

1. IS SOCIAL LOAFING A UNIVERSAL EFFECT WITH DISASTROUS CONSEQUENCES FOR GROUP PERFORMANCE?

Collective work and cooperation among individuals are fundamental to the establishment of interpersonal relations and bonds, the creation and functioning of organizations, and the development of society. Throughout history, the great achievements of humankind have been the result of this kind of cooperative action. Intuitively, it is assumed that working with others in a group has an energizing effect, causing people to work hard. In addition, the beneficial effects of groups are generally acknowledged (e.g., "unity is strength," "many hands make light the work"), and collective efforts are rewarded. The performance of people working together in teams can be very high if the people are ready to cooperate with each other.

Nevertheless, people often witness the contrary, as was revealed by the early research of social psychology on group performance. People carrying out collective work are not always willing either to cooperate or to work hard. Under such conditions, they try to go unnoticed, hiding in the group or crowd and avoiding contributing fully to the collective product. At the same time, they are suspicious of and distrust their fellow group members and worry that others might not make their respective contributions.

One important research tradition confirms that people tend to make less effort when their contributions are combined with those of other people, in contrast to the effort and interest they show when performing individually. In such circumstances, it may be observed that "the whole is less than the sum of its parts." This effect, known as social loafing, is related to two other effects, free riding and the sucker effect, and has been found to be sufficiently robust and generalized among the different types of tasks and populations to be considered universal, that is, an inherent feature of group performance. If this is the case, concern is warranted because groups are an ever-present feature of organizations and society. Most work is carried out in groups, and a large part of this work involves performing tasks in which the respective contributions of all members are combined. Is social loafing a universal effect, and does it have disastrous implications for group performance?

To address these issues, verify the extent of social loafing, identify how it works and what consequences it has, and find out how to prevent it or counteract its effects, this article analyzes this phenomenon throughout the historical evolution of the research carried out and the main theoretical models used to explain it. Finally, it presents practical suggestions for preventing and controlling social loafing.

2. A BRIEF HISTORICAL OVERVIEW

Scientific interest in social loafing is relatively recent. In 1913, Ringelmann performed an experiment that consisted of asking volunteers to pull on a rope as hard as possible, both individually and in groups of various sizes. He discovered a nearly linear reduction in the amount of strength exerted in groups compared with expectations based on individual performances. Although average group productivity increased, the result fell short of the maximum predicted as new members were added to the group. These results drew attention to the possibility that the performance of collective tasks could have a demotivating effect on group members. Initially, Ringelmann surmised that this could be due to coordination problems. Thus, the individuals might be having difficulty in combining their efforts in a way that was fully effective, and these difficulties would be more likely to emerge in large groups. A well-known "law" established by Steiner in 1972 states that a group's real productivity is equal to its potential productivity minus losses due to defective processes such as coordination.

Ringelmann's experiment was not replicated until 1974, when a number of inquiries into the causes of social loafing were undertaken. Some ingenious experimental strategies were designed in an attempt to reject the original explanation of coordination-related losses. In the replication of Ringelmann's experiment carried out by Ingham and colleagues in 1974, the participants were blindfolded to prevent distraction and placed at the head of a group with instructions to pull on a rope. However, the rest of the group members were actually the experimenter's confederates, and they did not really pull on the rope. In fact, the "group" was fictitious and had only one member who believed that he or she was part of a team. The outcome observed by Ringelmann was also obtained in this situation. That is, as the

apparent size of the group increased, performance dropped, confirming that the effect was not due to coordination difficulties; rather, it was due to motivation loss related to the size of the group.

Latané and colleagues obtained similar results in 1979 when they replicated the experiment with a different task. Blindfolded participants who also wore earplugs were asked to cheer or clap as loud as they could. Again, motivation decreased as the apparent size of the group increased. Participants seemed less willing to make an effort if they felt that others might not be doing so. The authors called this reduction in the effort of members when working in a group social loafing, and they claimed that the phenomenon was widespread—a kind of "social disease" with negative consequences for individuals, institutions, and society. By way of example, they mentioned the former Soviet Union's collective farming systems.

Subsequent research confirmed the consistency and general prevalence not only of social loafing but also of two other closely related effects: free riding and the sucker effect. Thus, it was shown that social loafing was not confined to physical tasks (e.g., rope pulling, cheering) but also occurred in other cognitive, creative, evaluative, and work-related tasks. Moreover, it affects both men and women, as well as people of all ages, and was to be found in both Western and Eastern cultures. (The experiments were replicated in countries such as India, Japan, and Taiwan.)

Contrary to the original supposition, however, research has revealed that there is nothing inherently demotivating in the group context. Thus, social loafing cannot be considered a feature of group performance versus individual performance because it arises only under certain specific conditions, whereas in many other situations working in groups actually increases motivation.

3. THEORETICAL EXPLANATIONS

A variety of theories have been proposed in an attempt to explain the social loafing effect. Although all of these have their limitations and are open to criticism, they have allowed researchers to progressively expand their knowledge of social loafing. A summary of the main theories follows.

3.1. Social Impact Theory

The social impact theory treats social influence as a function of the intensity, immediacy, and number of sources and targets present. When a source (a person) insists on hard work during the performance of a collective task, demands from the source are diluted among the various targets (all group members) and, therefore, have little impact. Effort drops. People feel less responsible for task performance as the size of the group increases.

3.2. Arousal Reduction

Another explanation, which builds on the social impact concept, is based on arousal reduction theory. Research on social facilitation indicates that arousal increases when a person performs a task in the presence of others (i.e., a coactive task). This causes people to improve their performance in simple or familiar tasks (dominant responses) but has the opposite effect when they are asked to carry out complex or unusual tasks. However, the situation is different with social loafing. People are performing collective tasks rather than coactive tasks, and in this case the participation of others reduces arousal because the group's members as a whole become the targets of external social influence. Thus, different results are obtained depending on the circumstances. In complex tasks, individuals might do better working collectively than coactively because they do not feel as though they are the exclusive targets of attention.

3.3. Identification and Evaluation Potential

Many interpretations of social loafing refer to the absence of identification and evaluation of individual effort in the performance of collective tasks. Because the input of each individual is combined with the inputs of others to produce a single group result, the contribution made by each individual becomes indistinguishable from the contributions made by other members of the group. As the size of the group increases, so does anonymity. This makes it easier for individuals to shirk—to "get lost in the crowd"—while nobody can blame them or make them responsible for underperformance. If, on the other hand, each person's input is identifiable (e.g., if a microphone is placed in front of each person who is asked to clap or cheer in the experiment described earlier), social loafing drops. However, it is not sufficient for the input simply to be identifiable. It must also be evaluated, not only or necessarily by an external observer or the group but also by the individuals themselves.

From this point of view, social facilitation and social loafing can be considered two opposite phenomena explained by the same process, that is, the concern with evaluation. This contributes to heightened performance not only when simple tasks are performed individually alongside coactors but also when complex tasks are addressed collectively.

3.4. Dispensability of Effort

An important exception occurs when individuals perform disjunctive tasks, requiring input from only one member of the group. Under these conditions, individuals who believe that their efforts do not affect final group performance, although the input may be perfectly identifiable and capable of evaluation, seem unwilling to make the effort required to perform the task because they believe that other members will achieve the same result with less effort. If the most capable or skilled individual shares this perception, he or she will probably be ready to compensate the others. Otherwise, the free riding effect (i.e., other members not doing their share) and the sucker effect (i.e., unwillingness to let other group members exploit oneself) are likely to emerge. However, if the task is conjunctive (i.e., all group members are needed to perform the task), the effort put in by the most skilled individual is likely to fall.

3.5. Self-Efficacy

The two opposite phenomena of social loafing and social facilitation can be explained through the self-efficacy theory. Thus, the effort that individuals are willing to make will depend on the level of self-efficacy for the task. If they perceive themselves as having high self-efficacy for the task, they will probably be willing to work hard when they perform alone or alongside coactors (social facilitation). Individuals in this situation feel responsible and self-confident about their performance. In contrast, individuals who believe that they have low self-efficacy for the task are less likely to work hard, especially when their performance is being evaluated (social loafing).

3.6. The Collective Effort Model

The collective effort model (CEM) is an integrative model that incorporates elements from a variety of theories, including the expectation–value models of effort, the social identity theory, and self-evaluation processes. The expectation–value theory states that individuals make more effort under certain circumstances such as when they believe that (a) they will achieve more if they work hard (expectation), (b) good performance will be rewarded (instrumentality), and (c) the rewards are valuable and desirable (value or importance).

CEM extends this theory to account for individual motivation in group settings. In collective tasks, effort depends on the perceived relationship among (a) individual and group performance, (b) group performance and results, and (c) group and individual results. According to the theory, individuals seek to maximize the expected utility of their actions. Therefore, they will make an effort in performing a collective task if they consider their efforts to be instrumental in obtaining results that they value personally. The model assumes that the links among the different variables are weaker when performing collective tasks, so that individuals' perceptions of the instrumentality of their efforts to achieve valued goals are more uncertain when working in groups.

The determining factor in the evaluation process is not so much the result itself as an individual's own evaluation based on certain objective and subjective criteria. Among the latter, one important criterion is the group to which an individual belongs and identifies. As social identity theory explains, to the extent that the individual identifies with the positive attributes and results of the group, he or she gets a positive self-evaluation that is incorporated into his or her personal identity.

A meta-analysis of 78 studies of social loafing conducted by Karau and Williams in 1993 confirmed that social loafing has a clear and widespread effect (it is present in all of the studies considered even though the studies are quite different in themselves) and that it increases with the group size. The results also found patterns that support the rationale and implications of CEM. On this basis, social loafing is less probable when individuals work in small groups rather than large groups, when they perform interesting tasks rather than irrelevant tasks, when they work with people they respect (as friends) rather than with strangers, when they perceive their contributions as essential to the final product rather than as redundant, and when they expect their companions to perform poorly rather than skillfully.

3.7. Other Integrative Models: Instrumentality, Value, and Equity

In addition to CEM, other comprehensive theoretical models that guide current research have been proposed.

For example, in 2003, Rutte included in his model the concept of equity as a mediating mechanism in addition to instrumentality and the value of the result. This corrected and refined some of the assumptions made in CEM. When individuals perform collective tasks, they judge their own contributions to the collective performance on the basis of the inputs made by other group members and they make an effort to bring about a fair exchange (i.e., balanced input–output relationship). The risk of free riding is always present in groups, although people dislike the thought because such behavior is unjustifiable. Consequently, they make less effort to avoid being suckered, and this can trigger a contagious spiral of demotivation among all members of the group. This integrative explanation combines two approaches derived from previous theories: the rational approach, from expectation–value theory, and the moral approach, from equity theory.

4. IMPLICATIONS OF SOCIAL LOAFING

Karau and Williams's meta-analysis revealed the generalization of social loafing that makes it an issue of great concern because many of the conditions in the studies reviewed are present in most settings in which people work collectively, both in business and in public institutions. In different surveys carried out in these settings, workers themselves identified social loafing as one of the main problems of teamwork. However, although everybody acknowledged the phenomenon, nobody personally admitted to loafing.

The social loafing effect is especially problematic in educational settings due to the special emphasis placed on training activities based on collaborative groups. Current studies of the phenomenon pay special attention to new ways of group working such as virtual work teams and groups mediated by information technologies.

There are a number of observations that might be made about the general effect identified in the meta-analysis. With regard to the studies analyzed, nearly all were laboratory experiments carried out with ad hoc groups created out of context and having neither a past nor a future. When working with more realistic groups, different results are obtained. For example, in 1991, Erez and Somech analyzed groups of people who had known each other for at least the previous 6 months working under differing conditions of realism (with or without specific goals, communication, reinforcement, and collective values). In this case, social loafing was the exception rather than the rule, occurring in only 1 of the 16 experimental situations.

The results of the meta-analysis contradicted the literature of organizational psychology in support of work teams. Currently, for example, major changes are being made to organizational structure to replace the organization of work around individual jobs with specifically designed team-based organizations, as described by Mohrman and colleagues in 1995. Only teams that integrate a variety of skills and bring together sufficient knowledge, expertise, and experience are capable of providing agile and flexible responses to contemporary organizations' demand for competitiveness and innovation. Groups are essential if organizations are to be made more competitive (e.g., by applying knowledge management strategies) because they are much more efficient learning systems (i.e., they learn more and better) than are isolated individuals.

Compared to individual work and competitive relationships, cooperation substantially improves productivity and strengthens interpersonal relationships. Cooperation involves understanding that individual success and group success are intertwined and that assistance, support, and integration actions will contribute to the team's success. Cooperation is the essence of work teams and provides the basic justification for the creation of links between the individual and the organization and between the organization and its setting. In 2003, these concepts were discussed in a handbook edited by West and colleagues.

Finally, there is a strong tradition of research in relation to group motivational gains, as described by Stroebe and colleagues in 1996. These gains arise for a number of reasons and are generated by the distribution of the relative strengths of the group that provide incentives that are not present in individual situations. These incentives are related to social compensation (increased motivation in the most skilled or strongest group member, who is aware that some less capable or weaker members will not be able to contribute to the group and is ready to make up for their weaknesses), interpersonal competition (based on the motive to be better than others, and should be evoked when it is possible to invert the existing rank order), and the Köhler effect (increased motivation in the weakest group members, who try to match their performance to that of a stronger companion).

Therefore, the relationship between individual productivity and group productivity presents evidence of both losses and gains. The losses are perhaps better known and, in any case, are more marked. However, their effect is limited to certain situations.

5. PREVENTION AND CONTROL OF SOCIAL LOAFING

The social loafing effect is not inevitable. In fact, there are a number of well-known factors that can prevent and control social loafing and can also improve motivation in groups. In general, the integrative models and research results show that it is necessary to promote situations that develop the expectation that effort is necessary to obtain results and increase their value and that such endeavors are guided by equitable criteria. Some of these factors are more directly related to the tasks that individuals must perform, others are related to group functioning, and still others depend on organizations. Nevertheless, it is hard to separate these factors in practice.

5.1. Task-Related Factors

Some procedures call for adequate design or reappraisal of the tasks assigned to the group members. Thus, boosting instrumentality and reinforcement requires designing tasks on the basis of the factors shown in Table I (e.g., meaningful tasks, intrinsically interesting or attractive tasks). Some of these strategies are in fact applied by the managers of certain kinds of teams such as elite sports teams, whose players are individually filmed and evaluated to identify their personal performance.

5.2. Group-Related Factors

Other factors are related to the composition and processes of the group as well as to its relationships with other groups. Table I reflects a number of significant group characteristics that help to prevent social loafing.

Certain competition- or high goal achievement-oriented groups (e.g., sports teams, elite combat units) promote specific strategies to achieve additional motivation, for example, establishing rules oriented toward maximum effort, encouraging strong feelings of group loyalty, and increasing arousal.

Team training and development programs are fundamental to promoting the majority of these factors. For example, training in processes and group skills, whether behavioral or cognitive (e.g., communication, coordination of effort, shared mental models, transactive memory systems) helps to develop the perception that individual efforts are inextricably linked to group performance, to raise awareness of the characteristics of group members, and to increase convergence within the group with regard to both tasks and goals.

The strategies usually employed in team development are likewise essential to the prevention of social loafing inasmuch as they allow for improving individuals' skills in four main areas that were broadly discussed by Salas and colleagues in 1999: (a) setting and achieving goals; (b) developing communication, support, and trust; (c) identifying problems and generating, applying, and evaluating solutions; and (d) improving comprehension of the responsibilities and obligations of other members' roles. Therefore, the strategies used in team development are very useful for improving members' motivation to the extent that personal responsibility is enhanced, task relevance is recognized, commitment among members and to the group's shared goals is promoted, and the group's social identity is developed.

The role played by the leaders or managers is crucial to the extent that they are capable of forming the group, designing the physical and organizational setting in which it operates, and influencing members' perceptions, expectations, and behaviors. For example, the leader can make the group see that members' skills are necessarily mixed and can develop the expectation that every individual's input is essential. In addition to directly influencing members' behavior through supervision, direction, and control, the leader can appeal to their values and ideals to achieve organizational goals and promote change and innovation (transformational leadership).

5.3. Factors Related to the Organization

Other factors are related to the organizational setting in which groups and individuals are embedded. The design of the task, the organizational structure (communication, power, and decision structure), human resources policies and practices, and the organizational culture may create specific patterns that directly prevent social loafing and lend coherence and support to the strategies referred to earlier (Table I).

With regard to incentives, organizations can establish a system of individual and collective external reinforcement and can contribute to the development of internal reinforcement (collective or individual).

TABLE I
Factors Reducing or Eliminating the Social Loafing Effect

Level of analysis	*Factors and strategies*
Tasks	• Task evaluation (by supervisors, colleagues or the individuals themselves) allows the reinforcement of actual behaviors.
	• Different or unique tasks ensure that group members perceive that their inputs to the collective product are not redundant but rather necessary and complementary.
	• Tasks involve incentives to contribute.
	• Tasks reduce the costs of contributing.
	• Tasks are intrinsically interesting or attractive.
	• Tasks produce personal involvement.
	• Tasks are meaningful.
	• The collective results achieved should be considered valid and important for the individuals and their reference groups.
Group	• The group should be small.
	• The group should be made up of individuals who share a relationship based on respect and trust.
	• Group members should be aware of each other's skills as well as of each other's difficulties with some tasks or the inability to carry them out.
	• The group should have high cohesion.
	• The group should clearly set performance goals (in addition to or instead of individual goals), and these goals should be highly valued.
	• The group should receive feedback and reinforcement (in addition to or instead of individual feedback).
	• Group members should consider group results to be valid and important.
	• There should be group norms to regulate the behavior patterns required of members to achieve goals.
	• Group members should share the same view of the situation and mental models.
	• Group members should share a perception of high collective efficacy (i.e., team potency).
	• Group members should value the group and identify with it, establishing a positive identity within the group and with regard to other collectives.
	• Team training and team development programs should be applied.
Organization	• Select group members in accordance with their skills and the tasks required.
	• Set goals that are both challenging and realistic.
	• Define standards or comparative criteria.
	• Provide nonevaluative conditions to test and use skills.
	• Establish conditions to use skills and perform tasks.
	• Provide regular feedback on task performance.
	• Provide the necessary resources and support actions.
	• Establish an incentive system.
	• Create a setting that leads to the development and strengthening of self-efficacy and self-esteem.

Shepperd reviewed these systems in a 1993 article. The most commonly applied method is to establish external reinforcements that are both collective (material or social) and individual. The organization can foster the development of individual and collective internal reinforcement. To enhance internal individual reinforcement, the organization can not only design interesting and meaningful tasks but also set specific performance goals to allow individuals to evaluate and compare their behavior on the basis of objective criteria. By creating teams, acknowledging their autonomy, empowering their members, and using cultural artifacts (e.g., symbols, myths), the organization can also contribute to the development of internal collective reinforcements so that group members identify with the teams and the organization, feel proud to belong to them and feel responsible to contribute to them, and feel capable of coping with shared goals and of accepting collective successes as personal triumphs.

Although the use of collective reinforcements is very important, the conditions under which they are effective are unclear, according to a 1998 review by DeMatteo and colleagues. This is because they depend both on the characteristics of the reinforcement itself (size, frequency, and allocation procedures) and on the characteristics of the individuals, groups, and organizations concerned.

According to experts in the field, both the reinforcement system and the mechanisms used to set goals and provide feedback require the organization to guarantee a work setting that ensures a good match between team members' and teams' motivation and the organization's own goals. Similarly, the organization should avoid any dissociation or conflict between motivational factors.

To conclude this review of the factors that prevent social loafing, it is necessary to at least mention individual variables related to personality traits that can give rise to major differences among the members of a team. Among these variables, the need to belong and feel affiliated, a collectivist orientation, and the acceptance of Protestant ethics have been identified as key to the prevention of social loafing.

The prevention of social loafing and the enhancement of team members' motivation are complex topics that can hardly be explored fully in this article. However, it can be noted that they depend on a combination of variables that, in turn, present their own peculiarities, including the following:

- The tasks should not be confined to simple additive operations; rather, they should demand the synchronized combination of inputs from a variety of specialists.
- The teams should not be fixed, durable, and real; rather, they should be temporary, requiring members' partial and multiple affiliations, and also should also present some degree of virtuality.
- Goals and tasks should not be set in isolation from one another; rather, they should present varying degrees of interdependence.
- Decisions should be made not only on the basis of rational criteria but also taking into account moral and emotional factors.
- The feeling of belonging to a group or an organization and involvement are not always based on economic transactions or regulated by formal contracts; rather, they can also depend on ideals and values as well as on some kind of commitment.
- Work relationships are not confined to bonds between the individual and other group members but also include associations with other groups and organizations.

These variables represent the main topics of current research into work teams in the organizational setting. Results from this research will provide a better understanding of the keys to the motivation losses and gains that arise when people work in teams. This could allow for the achievement of maximum performance situations, where the individual's inputs are recognized and combined efficiently with the contributions made by other members of the team and organization while avoiding the temptation to overvalue the individual's contributions but without annihilating them in favor of a tyranny of collective values in this age of "groupism."

See Also the Following Articles

Groups, Productivity within ■ Groupthink ■ Intergroup Relations and Culture

Further Reading

Baron, R. S., Kerr, N., & Miller, N. (1992). *Group process, group decision, group action*. Buckingham, UK: Open University Press.

DeMatteo, J. S., Eby, L. T., & Sundstrom, E. (1998). Team-based rewards: Current empirical evidence and directions for future research. In B. M. Staw, & L. L. Cummings (Eds.), *Research in organizational behavior* (Vol. 20, pp. 141–183). Greenwich, CT: JAI.

Erez, M., & Somech, A. (1996). Is group productivity loss the rule or the exception? Effects of culture and group-based motivation. *Academy of Management Journal*, *39*, 1513–1537.

Ingham, A. G., Levinger, G., Graves, J., & Peckham, V. (1974). The Ringelmann effect: Studies of group size and group performance. *Journal of Experimental Social Psychology*, *10*, 371–384.

Karau, S. J., & Williams, K. D. (1993). Social loafing: A meta-analytic review and theoretical integration. *Journal of Personality and Social Psychology*, *65*, 681–706.

Karau, S. J., & Williams, K. D. (2001). Understanding individual motivation in groups: The collective effort model. In M. E. Turner (Ed.), *Groups at work: Theory and research* (pp. 113–141). Mahwah, NJ: Lawrence Erlbaum.

Kozlowski, S. W. J., & Bell, B. S. (2003). Work groups and teams in organizations. In W. C. Borman, D. R. Ilgen, & R. J. Klimoski (Eds.), *Industrial and organizational psychology* (Vol. 12, pp. 333–375). New York: John Wiley.

Latané, B., Williams, K., & Harkins, S. (1979). Many hands make light the work: The causes and consequences of social loafing. *Journal of Personality and Social Psychology*, *37*, 822–832.

Locke, E. A., Tirnauer, D., Roberson, Q., Goldman, B., Latham, M. E., & Weldon, E. (2001). The importance of

the individual in an age of groupism. In M. E. Turner (Ed.), *Groups at work: Theory and research* (pp. 501–528). Mahwah, NJ: Lawrence Erlbaum.

Mohrman, S. A., Cohen, S. G., & Mohrman, A. M. (1995). *Designing team-based organizations: New forms for knowledge work*. San Francisco: Jossey–Bass.

Rutte, C. G. (2003). Social loafing in teams. In M. A. West, D. Tjosvold, & K. G. Smith (Eds.), *International handbook of organizational teamwork and cooperative working* (pp. 361–378). Chichester, UK: Wiley.

Salas, E., Rozell, D., Driskell, J. D., & Mullen, B. (1999). The effects of team building on performance: An integration. *Small Group Research*, *30*, 309–329.

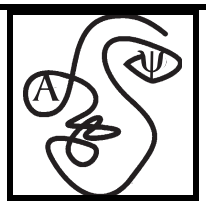
Shepperd, J. A. (1993). Productivity loss in performance groups: A motivation analysis. *Psychological Bulletin*, *113*, 67–81.

Steiner, I. D. (1972). *Group process and productivity*. New York: Academic Press.

Stroebe, W., Diehl, M., & Abakoumkin, G. (1996). Social compensation and the Köhler effect: Toward a theoretical explanation of motivation gains in group productivity. In E. Witte, & J. H. Davis (Eds.), *Understanding group behavior: Small group processes and interpersonal relations* (Vol. 2, pp. 37–65). Mahwah, NJ: Lawrence Erlbaum.

Weaver, J. L., Bowers, C. A., Salas, E., & Cannon-Bowers, J. A. (1997). Motivation in groups. In M. M. Beyerlein, & D. A. Johnson (Eds.), *Advances in interdisciplinary studies of work teams* (Vol. 4, pp. 167–191). Greenwich, CT: JAI.

West, M. A., Tjosvold, D., & Smith, K. G. (Eds.). (2003). *International handbook of organizational teamwork and cooperative working.* Chichester, UK: Wiley.



Social Networks

Mieneke W. H. Weenig
Leiden University, Leiden, The Netherlands

1. Introduction
2. Social Networks
3. Major Research Findings
4. Conclusions
Further Reading

GLOSSARY

bridge A person who links two or more cliques in the network from his or her position as a member of one of the cliques.
centralization The extent to which ties in a network are directed to a central individual in the network.
cliques Subsystems in the larger network whose elements interact with each other relatively more frequently than with other members of the network.
connectedness The extent to which individuals in a network are linked by direct or indirect ties.
density Ratio of the number of existing ties to the number of possible ties.
homophily The degree to which two interacting individuals are similar in certain attributes such as beliefs, norms and values, and socioeconomic status.
interlocking network A network in which all members communicate with each other.
knots (or stars) Individuals who have relatively many links with other members of the network or clique; they are central persons in the network like spiders in a web.
liaison A person who links two or more cliques in the network but who is not a member of any clique.
multiplexity The degree to which a network link is used for various topics.
network structure The pattern of communication links in a system.
sociometry A means of obtaining quantitative data about who communicates with whom in a system.
structural equivalence The degree to which two individuals are linked to the same set of other persons but not necessarily to each other.
tie (or link) A communication relationship between two units (usually individuals) in a system.
tie strength A combination of the amount of time invested, the emotional intensity and intimacy, and the reciprocal services that characterize a tie.

A social (or communication) network is an aggregation of individuals who are tied to each other by patterned flows of information. Ties can vary in strength, depending on the frequency of use and on the degree to which the ties are used for multiple message contents. Weak ties are especially important for the diffusion of information, whereas strong ties are more important with respect to the flow of influence.

1. INTRODUCTION

Humans are social beings, and the social context plays an important role in what they do (behavior) and what they think (attitudes and beliefs). For instance, when an individual wants to buy a new car, he or she would probably base the ultimate decision on a mixture of three kinds of information: (a) the individual's own

©2004 Elsevier Inc.
All rights reserved.

experiences with cars; (b) information obtained from brochures, television, and magazines; and (c) information obtained from other persons such as experiences or advice from personal friends and acquaintances and advice from experts (e.g., the car salesperson). In addition, the individual would observe what cars his or her friends and neighbors have bought themselves and probably would not buy a car that he or she expect the social environment to reject.

In other words, individuals base their purchase decisions at least partly on information, expertise, and expectations from others that is communicated mainly through interpersonal networks. Thus, it is very important to know something about the social context of a target group and how it affects people to develop and implement more effective promotion (or "persuasion") campaigns.

2. SOCIAL NETWORKS

The social context of a person, group, or community can be conceived as an aggregation of individuals. Some of them communicate with each other, whereas others do not. The pattern of all such communication ties is called a communication network or social network. Basically, such a network indicates who communicates with whom. Together, these "communication links" constitute a network of direct and indirect ties between people. The communication link or tie is the basic unit of each communication network. It is the dyadic link of interpersonal communication between two individuals.

Ties may vary in strength, depending on the frequency of use and on the degree to which a tie is used for multiple message contents (also called "multiplexity"). In 1973, Granovetter defined the strength of ties as "a (probably linear) combination of the amount of time invested, the emotional intensity and intimacy (mutual confiding), and the reciprocal services which characterize the tie." Thus, close relatives and friendship relations can be regarded as strong ties.

The formation of a strong tie between two network members is more likely in case of homogeneity with respect to norms, values, important beliefs, socioeconomic status, and demographic characteristics. This is referred to as the "homophily principle." In general, the more homogenous two individuals are, the higher the probability of a strong tie between them. However, in the immediate geographic environment, spatial distance appears to be a decisive factor. Therefore, in neighborhoods, the probability of a rather strong (at least in terms of high frequency) tie between next-door neighbors is relatively high and less dependent on similarity with respect to social and demographic characteristics.

Apart from the strength of ties, the most frequently used characteristics of ties in social network research are the stability (duration over time) and direction of ties (unidirectional vs bidirectional). In addition, several structural characteristics of social networks have been identified, including the existence of cliques, knots, bridges, and liaisons. Finally, various network indexes have been developed for network analyses, including the degree of density, connectedness, and centralization.

3. MAJOR RESEARCH FINDINGS

3.1. Efficiency of Information Diffusion in Various Types of Networks

Some of the first studies on communication networks were performed during the 1950s by Massachusetts Institute of Technology researchers who were mainly interested in the relationship between network structure and efficiency of information diffusion. Basically, they used an experimental design in which participants could communicate with one another through various kinds of network structures, including a wheel, a chain, a Y, a circle, and a "comcon" (completely connected). Of these network forms, the wheel (where all information goes to one central person) was the most centralized network form, whereas the comcon (where all individuals can communicate with one another) was the least centralized network form. Results suggested that a centralized network is more efficient than a decentralized network, but later studies qualified this rule to apply mostly to relatively simple information and routine tasks.

The information diffusion process is also affected by the strength of ties. In his seminal article, Granovetter argued that weak ties, although used less frequently than strong ties, may nevertheless be more valuable for the diffusion of information than are strong ties because weak ties are more likely than strong ties to serve as bridges between subgroups (or "cliques") in the larger network. Thus, bridges to other cliques prevent information from remaining inside clique boundaries and bring new information to a clique, a notion that has received support from several investigations. However, within a clique, strong ties seem to be more important for information diffusion.

3.2. Network Influences on the Diffusion of Innovations

Social network influences on people's attitudes and behavior have been investigated most extensively in the areas of the diffusion of innovations, marketing, and organizational communication. Research on the diffusion of innovations has addressed the question of which factors determine people's decisions to adopt innovations and how these adoption decisions spread in the social system. Rogers defined an innovation as "an idea, practice, or object that is perceived as new by an individual or other unit of adoption." Innovation diffusion is the process by which the adoption of an innovation spreads over time in a social system through communication channels among its members. The adoption rate is the relative speed with which an innovation is adopted by all members of a social system.

Rogers identified "opinion leaders" as crucial members of a social system. An opinion leader is defined as an individual who "is able to influence other individuals' attitudes or overt behavior informally in a desired way with relative frequency." According to Rogers, major characteristics that distinguish opinion leaders from their followers include greater exposure to mass media and a more "cosmopolitan" orientation, more personal ties within the community as well as with people beyond the group's boundaries (including change agents), and a higher social status in comparison with their followers. Computer simulations have shown that diffusion occurs up to three times faster when it is initiated by opinion leaders than when it is initiated by random network members.

One of the most classic innovation diffusion studies is Coleman and colleagues' 1957 study on the diffusion of a new medical drug among a group of doctors. In this study, socially integrated doctors, described as doctors who received three or more sociometric choices as "friends" by their colleagues and thus were considered to be "opinion leaders" of the network, were observed to have adopted a new drug much earlier than did less integrated doctors (who were mentioned as friends less frequently). The moment that nearly all of the opinion leaders had adopted the new drug, the diffusion rate of the opinion leaders' followers accelerated rapidly. This was interpreted as the result of interpersonal communication about the new drug. However, more recently, Burt reanalyzed Coleman and colleagues' data and demonstrated that the results did not reflect interpersonal conversations among colleagues (so-called "contagion by cohesion"). Instead, Burt found that the data reflected a (nonverbal) process of social comparison and competition among doctors who occupied structural equivalent positions in the communication network of the doctors (i.e., had similar patterns of ties within the network), irrespective of whether they had direct ties with each other (so-called "contagion by structural equivalence"). Based on these data and on data from business and politics, Burt showed that the strength of relationships (i.e., cohesion) increases the adoption rate only in the case of weakly equivalent individuals. According to Burt, contagion occurs regardless of cohesion for strongly equivalent people, whereas contagion does not take place at all for nonequivalent people.

Although contagion might not necessarily require the existence of direct ties but rather might require an equivalent position in the network, some studies indicate that contagion does take place at least partly through the influence of direct ties. An example is a 1981 study by Rogers and Kincaid on the diffusion of family planning measures in Korea. It showed that individuals were more likely to adopt the innovation if a larger proportion of the persons with whom they most often talked about family planning consisted of individuals who had already adopted the innovation.

3.3. The Impact of Word-of-Mouth Communication

Other indications of the influence of direct communication links on adoption decisions come from consumer and marketing studies on the occurrence and impact of "word-of-mouth" (WOM) recommendations. A large number of studies on the antecedents of purchase decisions have shown that substantive percentages (sometimes even large majorities) of buyers mention WOM recommendations as important or major sources of information or influence concerning the purchase of a specific product or service. In fact, the frequency of occurrence and the apparent influence of WOM communication have even encouraged several manufacturers to implement promotion strategies of their products that are based nearly exclusively on stimulating WOM recommendations within their target groups by means of strategically providing free samples of their products (e.g., Häagen-Dazs ice cream). However, in spite of the acknowledged importance of WOM recommendations for purchase decisions by manufacturers and scientists alike, surprisingly little research has been performed on the

dynamics and antecedents of WOM recommendations. One of the most prominent exceptions is Brown and Reingen's 1987 research. These authors found that weak ties served as information bridges through which WOM information traveled from one subgroup in the network to another, whereas strong and homophilous ties were more activated for the flow of referral information and also were perceived as being more influential than weak ties.

Research on negative WOM communication is even scarcer, but findings suggest that, in general, negative WOM communication has a stronger influence on consumers' evaluations of brands and products than does positive WOM communication. This is consistent with the so-called "negativity effect" that is often found in the impression formation and decision-making literature. It means that negative information is given greater weight in judgments than is positive information of equal extremity. The most accepted explanation for the negativity effect is that negative information is generally perceived as more diagnostic than positive information for categorizing target objects into evaluative categories.

In 2002, Ahluwalia suggested that the influence of negative WOM communication on product evaluations depends on the causal attributions that recipients make after hearing negative WOM messages. For instance, if a recipient infers from a message that only one person has had a negative experience with a particular brand or product, whereas most people have had positive experiences with that product (low consensus), or when a recipient perceives the communicator to be a person who has negative opinions about many products or brands (low distinctiveness), the negative message is attributed to (i.e., is perceived as being caused by) the communicator rather than to the product to which the negative WOM message refers. In that case, product evaluations might even increase due to negative WOM communication.

Other recent research performed by Laczniak and colleagues suggested that the negativity effect applies mostly to consumers who are unfamiliar with a specific product or brand but who are motivated to evaluate that product as accurately as possible with an "open mind" to all kinds of information. On the other hand, if consumers are committed to a particular product or brand, they usually want to defend their choice. In that case, a positivity effect is more likely than a negativity effect, meaning that more weight is given to positive information that is consistent with consumers' existing (positive) opinion concerning the product than to negative (i.e., inconsistent) information.

In sum, the stronger influence of negative WOM communication on product evaluations seems to be restricted to situations where recipients are relatively unfamiliar with the product to which the negative message refers, attribute the negative information to the product and not to the communicator, and are motivated to make an accurate evaluation of the product.

WOM communication generally seems to occur more frequently among strong ties than among weak ties and with more impact on attitudes and behavior. However, whenever WOM communication occurs among weak ties, it seems to be evaluated more often as new and surprising information.

Apart from direct communication links, indirect communication links also may play an important role in the diffusion process. Several rumor studies have shown that secondhand information (i.e., information transmitted by indirect ties) is often more polarized (positively or negatively) or exaggerated than is firsthand information. Thus, secondhand information about an individual's positive or negative experiences with an innovation may be more "extreme" than the original information.

3.4. Network Influences on the Diffusion of Information Versus Adoption Decisions

Before an individual can make a decision on whether to buy a product or adopt an innovation, he or she should first be informed or made aware of its existence and also learn something about its features or characteristics. However, in many studies, no distinction is made between the process of information diffusion and social influence on adoption decisions. Such a distinction was made in a study by Weenig and Midden in 1991. They investigated the influence of a target group's existing social network on the information diffusion process and effects (i.e., positive adoption decisions) of several small-scale communication programs that aimed to promote the adoption of energy conservation measures. In each case, the program was developed and implemented by a project group that consisted of one or two professional change agents and several volunteers from the target group. In all cases, program awareness was positively related to the number of weak ties that residents had within the neighborhood and was unrelated to the number of strong ties that they had within the neighborhood, thereby supporting Granovetter's hypothesis on the strength of weak ties. Program awareness and attention to program activities

were also higher among residents who were personally acquainted with at least one of the volunteers from the project group than among residents who had no direct ties to the project group. The presence of direct ties to the project group did not affect program effects (i.e., adoption decisions). Instead, adoption decisions were positively related to advice received from strong ties (i.e., friends or family members) and were unrelated to advice received from weak ties. This means that the strength of ties is important with respect to influence on adoption decisions. Brown and Reingen found similar results concerning WOM referral behavior (described earlier). Thus, weak ties are important for the diffusion of information, whereas strong ties are more important with respect to the flow of influence.

In one of the target groups, however, the adoption of cavity wall insulation appeared to be negatively related to the number of weak ties in the neighborhood. It appeared that this negative relationship reflected the awareness of a negative rumor—contending that cavity wall insulation would lead to moldy walls—that had been disseminated among residents. The negative relation between adoption of cavity wall insulation and number of weak ties means that awareness of this negative rumor occurred mostly through weak ties. This awareness resulted (both directly and indirectly) in fewer adoptions of cavity wall insulation—directly if the rumor was heard from more than one weak tie (consensus information) and indirectly if the rumor was heard from only one weak tie (when it triggered asking additional advice from strong ties).

4. CONCLUSIONS

From the research described in the previous paragraph, three important conclusions can be drawn. First, the quantity of existing communication ties in the social network of a target group affects the amount and speed of diffusion of information. That is, the larger the degree of connectedness among target group members, the more alternative routes will be available for the diffusion of information. The flow of information will usually be fastest among strong ties because they are (by definition) used more often than are weak ties. However, because strong ties tend to cluster into interlocking subgroups (cliques) in the larger network, new information from outside a clique is likely to enter a subgroup through a bridge or liaison, and such ties tend to be weak.

Second, the strength (or quality) of ties determines the potential of social influence that may occur within the target group. That is, people appear to be influenced most by their friends and close relatives or by other "near peers." The potential of social influence from friends and close relatives has been stressed and demonstrated in several social psychological research traditions, including Bandura's social learning theory (i.e., attractive persons are the most influential models), persuasion theory (i.e., friends are usually more persuasive than strangers), Festinger's social comparison theory (i.e., similar persons serve as reference persons more often than do dissimilar persons), and conformity theory (i.e., more compliance to group norms occurs in cohesive or tightly knit groups). In fact, it might well be that the influence that strong ties exert is a mixture of all these kinds of social influence.

Third, opinion leaders, so far as they can be identified in a network or subgroup, may accelerate the speed of information diffusion and adoption decisions to a considerable extent, mainly as a result of their relative centrality in the network in combination with their ties with people outside the (sub)group's network. Unfortunately, such opinion leaders do not exist in all communities. However, several studies have shown that opinion leaders can be "created," to some extent, by means of training volunteers from the target group to become experts on the relevant topic. Such volunteers are able to reach and stimulate a large proportion of the target group provided they have many personal ties with target group members and have at least some interest in the topic. This means that, apart from at least some interest in the relevant topic, the number of ties within the target group should be an important selection criterion with respect to the recruitment of volunteers. These volunteers must subsequently be trained to become paraprofessional change agents. The duration and intensity of this training depend on the difficulty and complexity of the innovation and on the educational background of the volunteers, among other factors. For simple products, it might be enough to train the volunteers in how to provide free samples of the product to strategic persons (e.g., knots, bridges) in the network, whereas for more complex innovations, elaborate and lengthy training on various aspects (e.g., knowledge about the innovations, group dynamics, and communication processes) might be necessary.

This implies that it is crucial to have at least some knowledge of the existing social network of the target group. No doubt, the most comprehensive way in which to obtain this is through gathering sociometric data on the existence of ties, their strength, and their structure. However, even in their most simple form, such analyses

are complex, time-consuming, and costly. Moreover, for practitioners, they often result in unnecessarily detailed data, for in many cases change agents will not be interested in detailed descriptions of who communicates how frequently with whom. Instead, change agents may be more interested in the quantity and quality of existing ties at an aggregated level and in whether opinion leaders and bridges or liaisons can be identified. Local authorities, such as schoolteachers, church leaders, police officers, and community workers, can often provide such information. In that manner, essential information about network characteristics can be obtained quickly and easily and can be used subsequently to implement more efficient promotion strategies.

See Also the Following Articles

Groups, Productivity within ■ Groupthink ■ Intergroup Relations and Culture

Further Reading

Ahluwalia, R. (2002). How prevalent is the negativity effect in consumer environments? *Journal of Consumer Research, 29*, 270–279.

Brown, J. J., & Reingen, P. H. (1987). Social ties and word-of-mouth referral behavior. *Journal of Consumer Research*, *14*, 350–362.

Burt, R. S. (1987). Social contagion and innovation: Cohesion versus structural equivalence. *American Journal of Sociology*, *92*, 1287–1335.

Burt, R. S. (1999). The social capital of opinion leaders. *Annals of the American Academy of Political and Social Science*, *566*, 37–54.

Granovetter, M. S. (1973). The strength of weak ties. *American Journal of Sociology*, *78*, 1360–1380.

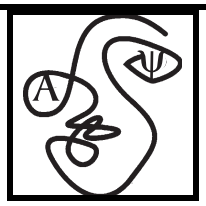
Laczniak, R. N., DeCarlo, T. E., & Ramaswami, S. N. (2001). Consumers' responses to negative word-of-mouth communication: An attribution theory perspective. *Journal of Consumer Psychology*, *11*, 57–73.

Rogers, E. M. (1995). *Diffusion of innovations* (4th ed.). New York: Free Press.

Rogers, E. M., & Kincaid, D. L. (1981). *Communication networks: Toward a new paradigm for research*. New York: Free Press.

Shaw, M. E. (1964). Communication networks. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 1, pp. 111–147). New York: Academic Press.

Weenig, W. H., & Midden, C. J. H. (1991). Communication network influences on information diffusion and persuasion. *Journal of Personality and Social Psychology*, *61*, 734–742.



Social Skills Training

Stephen N. Elliott and Sylvia C. Lang
University of Wisconsin, Madison, Wisconsin, USA

1. Introduction
2. Identification of Children in Need of Social Skills Training
3. Social Skills Training Procedures
4. Effectiveness of Social Skills Interventions
5. Selection of Interventions for Students with Different Problems
Further Reading

GLOSSARY

behavioral rehearsal A process in which the intervention target rehearses a specific social behavior and receives informational feedback on the rehearsal.

coaching A direct verbal instructional technique in which a "coach" (intervention agent) provides the intervention target with a standard for a specific behavior, the opportunity to rehearse that behavior, and informational feedback on the rehearsal.

discrimination training The process of training an individual to differentiate between situations and/or settings in which specific behaviors or actions would or would not be appropriate.

modeling A process in which the intervention target observes the social behavior(s) of a model for purposes of the intervention target acquiring new, socially appropriate behaviors.

social skills acquisition problem Characterizes children who have not acquired the necessary social skills to interact appropriately with others, or children who have not acquired a critical step in the performance of a given skill.

social skills performance problem Characterizes children who have appropriate social skills in their behavior repertoires but do not perform the behavior at acceptable levels and/or at appropriate times or settings.

Social skills interventions emphasize the acquisition and performance of prosocial behaviors and typically utilize nonaversive methods to teach these prosocial behaviors. Some social skills interventions, however, must concurrently focus on the reduction of interfering problem behaviors. Thus, social skills training has four primary objectives: promoting social skills acquisition, enhancing social skills performance, reducing or removing interfering problem behaviors, and facilitating the generalization and maintenance of social skills. Common social skills training tactics can be characterized as therapist directed, therapist and peer directed, or peer directed. These procedures can be further categorized into three theoretical approaches: operant, social learning, and cognitive–behavioral. Research indicates that social skills training overall is a moderately effective procedure for improving the social functioning of children age 3 to adults age 21.

1. INTRODUCTION

Behaviors such as sharing, helping, initiating relationships, requesting help from others, and ignoring teasing or aggressive reactions from others are examples of socially appropriate behaviors that enable children young and old to successfully function in social settings. Developing these types of social skills is one of the most important accomplishments of childhood and can lead to successful social relationships and academic functioning.

©2004 Elsevier Inc.
All rights reserved.

Social skills may be defined as socially acceptable learned behaviors that enable a person to interact with others in ways that elicit positive responses and to avoid negative responses. When an individual enacts social skills successfully, the results generally are positive perceptions of that individual's social competence. The acronym of CARES was developed by Gresham and Elliott in 1990 to characterize major clusters of social skills frequently targeted for intervention. These skill clusters are cooperation, assertion, responsibility, empathy, and self-control. Guevremont developed an alternative, but equally appealing, situational approach to characterizing clusters of social skills. Specifically, his skill clusters are social entry skills, conversational skills, conflict-resolution skills, problem-solving skills, and anger-control skills. Regardless of the terminology or approach used to characterize social skills, the behaviors of interest to most individuals conducting social skills training are those observable nonverbal and verbal interpersonal skills that maximize social engagement and social reinforcement.

The following are key assumptions to the assessment and treatment of social skills: (i) Social skills are primarily acquired through learning that involves observation, modeling, rehearsal, and feedback; (ii) they include specific, discrete verbal and nonverbal behaviors; (iii) they require specific behaviors that are effective in initiating, maintaining, and terminating interpersonal interactions and relationships; (iv) they are interactive by nature and entail effective and appropriate behavioral performances; and (v) they are situationally specific behaviors and are influenced by the characteristics, demands, and expectations operating in specific environments. Collectively, these five assumptions stress the multidimensional (verbal–nonverbal and initiating–responding) interactive, situation-specific nature of social skills. As such, effective interventions need to address (i) target behaviors involving both verbal and nonverbal communications used to initiate or respond to others and (ii) the ability to discriminate when to exhibit certain social skills given changing social situations.

2. IDENTIFICATION OF CHILDREN IN NEED OF SOCIAL SKILLS TRAINING

A number of methods, including rating scales, checklists, and sociometric nomination techniques, have been designed to identify children at risk for behavior problems. In general, social skills assessments have one of two purposes: identification/classification or intervention/program planning. A standard battery of tests or methods for assessing social skills does not exist. To increase the likelihood of accurate identification/classification decisions, it is recommended that one use direct observations of the target child and nontarget peers in multiple settings; behavioral interviews with the referral source and possibly the target child; rating scale data, preferably norm-referenced, from both a social skills scale and a problem behavior scale completed by more than one source; and sociometric data from the target child's classmates. Regarding intervention decisions, data contributing to a functional analysis of important social behaviors are imperative. These types of data usually result from multiple direct observations across settings, behavioral role-plays with the target child, and teacher and parent/guardian ratings of socially valid molecular behaviors.

Children identified for social skills intervention historically have been either individuals who have been rejected by their peers because they behave aggressively toward others or students with significant cognitive disabilities. During the past two decades, a large number of students with learning disabilities have also been targeted for social skills interventions. A number of researchers have documented that mildly disabled students who are classified as behavior disordered, learned disabled, or cognitively impaired frequently exhibit a significant number of social skills deficits and could benefit from individual or group social skills interventions. Social skills have also been documented to be an important part of academic success for nondisabled students, and thus the number of classwide school skills programs has also increased substantially during the past decade. The intervention strategies used for individual students are directly applicable, with only minor changes, to small groups of students, a classroom, or an entire school.

3. SOCIAL SKILLS TRAINING PROCEDURES

Social skills training procedures focus on positive behaviors and often use nonaversive methods (e.g., modeling, coaching, and reinforcement) to improve children's behavior. Therefore, use of these methods may enhance treatment acceptability and integrity and they can easily be built into the existing structure of a classroom or home environment. Social skills

interventions can be used with individuals or groups of students, and because they primarily concern increasing prosocial behaviors, all students can participate and potentially benefit from the interventions.

A large number of intervention procedures have been identified as potentially effective for social skills training with children age 3 to adults age 21. These procedures can be classified into three approaches that highlight common treatment features and assumptions about how social behavior is learned: operant, social learning, and cognitive–behavioral. In practice, many researchers and practitioners have used procedures that represent combinations of these basic approaches. However, these three approaches to intervention are used in this article to describe the basic procedures and to organize a review of their effectiveness.

3.1. Operant Intervention Procedures

Operant intervention procedures focus on discrete, observable behavior and the antecedent and consequent events that maintain that behavior. Behavioral control is achieved most often through the use of reinforcement and/or punishment paradigms that are contingent upon a specified behavior.

Children often fail to interact successfully because of a nonresponsive or inhospitable social environment. Here, the use of antecedent control can serve to facilitate a positive environment and elicit positive social interactions. With these procedures, a teacher or other intervention agent makes wide use of cueing and prompting. The key assumption with antecedent control procedures is that a child possesses the necessary social skills but he or she is not performing them at an acceptable rate. Two of the most frequently used antecedent strategies are peer social initiation and cooperative learning. Peer social initiation entails training similar-age peer confederates to initiate and maintain social interactions with a withdrawn or isolated child. Overall, this procedure has been effective in increasing positive social behavior in withdrawn children. It should be noted, however, that when a child's interaction rate approaches zero, peer initiations are likely to be less effective. Cooperative learning, on the other hand, provides a prosocial environment in which children work together to complete specified activities. This procedure requires children to cooperate, share, and assist each other to complete a task and, as such, represents an effective method for increasing the likelihood of positive social behaviors.

Many operant learning procedures, which involve the manipulation of antecedents and consequents, have been used to decrease interfering problem behaviors while simultaneously increasing positive social behaviors. These procedures are based on the premise that behaviors are maintained by functional relationships. In other words, reinforcement contingencies (positive or negative) perpetuate low rates of positive social interactions and high rates of negative social interactions. For this reason, it is assumed that a child can perform positive social interaction skills but does not demonstrate the skills due to lack of reinforcement. Two of the most frequently used methods to decrease inappropriate social behaviors are contingent social reinforcement and differential reinforcement. Contingent social reinforcement involves publicly reinforcing a child for socially appropriate behaviors. Although this procedure has been found to be effective, it can require an extraordinary amount of time to ensure that the reinforcement is delivered on a consistent basis. Unless reinforcement is given in a consistent manner, it is doubtful the intervention will be as effective. This type of contingent social reinforcement perhaps is best used to maintain social interaction rates established through the use of other social skills interventions. Other operant learning procedures used to modify social skills are differential reinforcement of other behavior (DRO), differential reinforcement of alternative behaviors (DRA), and differential reinforcement of low rates of responding (DRL). DRO and DRA are based on the omission of a specific behavior rather than the commission. Therefore, after a certain amount of time, reinforcement is given after any behavior except the target behavior.

DRL involves reinforcement for reduction in the frequency of a target behavior. Reinforcers may be delivered either for reduction in overall frequency of a response within a particular time period or for increased elapsed time between responses (interresponse time). DRO and DRL procedures are probably most effective, however, when used in conjunction with other social skills interventions, whereby one decreases socially inappropriate behavior while concurrently teaching or reinforcing positive social behaviors.

3.2. Social Learning Intervention Procedures

Social learning procedures can be traced back to the social learning theory of Bandura and Walters

(published in 1963) and Bandura (published in 1977). From this perspective, social behavior is the result of two types of learning: observational learning and reinforced learning. Social learning theorists differentiate between the learning and the performance of a response. This distinction enables social learning theorists to advocate the process of modeling as a means to acquire new, socially appropriate behaviors. Modeling also affects previously learned responses through its disinhibitory and/or cueing effects. Therefore, children are vicariously reinforced (positively or negatively) by observing a model receiving reinforcement for a behavioral performance. In essence, observers tend to inhibit responses that they see punished in others, whereas they are likely to perform modeled behaviors that elicit desired reinforcers. Modeling has substantial empirical support for the promotion of social skills with children and youth. A powerful development in the realm of social learning has been peer-mediated interventions. These interventions are based on the premise that peers can be effective change agents for other children with social skills performance deficits. Empirically, it has been shown that peers can serve to differentially reinforce appropriate social interactions and influence the occurrence of positive social behaviors. Peer-mediated interventions may be more successful than teacher-mediated interventions because peer confederates may be better able to consistently monitor and differentially reinforce their peers. Additionally, peer mediation can be cost-effective because it minimizes teacher involvement.

3.3. Cognitive–Behavioral Intervention Procedures

Cognitive–behavioral intervention procedures are a loosely bound group of procedures that focus on a child's internal regulation of his or her behavior. In particular, cognitive–behavioral approaches to social skills training emphasize a child's ability to problem solve and to self-regulate behavior. Two of the most frequently used cognitive–behavioral social skills procedures are coaching and problem solving. Coaching, unlike modeling, is a direct verbal instruction technique that involves a "coach" (usually a teacher or a psychologist, and occasionally a peer) who has knowledge about how to enact a desired behavior. Most coaching procedures involve three basic steps: (i) A student is presented with rules for, or the standards of, a specific behavior; (ii) a selected social skill(s) is rehearsed with the coach; and (iii) the coach provides specific informational feedback during behavioral rehearsal and gives suggestions for future performances. In some instances, modeling may be included in the coaching procedure to give the child a better understanding of the topography of the social behavior and, if praise is given, reinforcement may occur. Therefore, coaching, although conceptualized as a procedure that requires the child's cognitive skills to translate instruction into desired behaviors, can be enhanced with behavioral and/or social learning procedures. Coaching has received considerable empirical support as a social skills training procedure.

Several interventions have been developed that stress teaching children the process of solving social or interpersonal problems as a way to facilitate socially appropriate behaviors. As Weissburg noted in 1985, some of these intervention programs, which are largely classroom based, have been called social problem solving (SPS) programs, whereas others have been called interpersonal cognitive problem solving (ICPS) programs. ICPS programs generally place greater emphasis on cognitions that parallel social problem situations and employ narrower training procedures than SPS approaches; however, both use a similar training sequence to help students identify and cope with social problems. Briefly, the steps can be described as follows: (i) Identify and define the problem, (ii) determine alternate ways of reacting to the problem, (iii) predict consequences for each alternative reaction, and (iv) select the "best" or most adaptive alternative. Social problem-solving methods can be used with individual children or with entire classrooms.

4. EFFECTIVENESS OF SOCIAL SKILLS INTERVENTIONS

The popularity and widespread use of social skills training procedures have resulted in several traditional and meta-analytic reviews of the effectiveness of these procedures with children. Schneider and Byrne, who conducted the first major meta-analysis of social skills training studies in 1985, reported social skills interventions were more effective for preschoolers and adolescents than elementary children. No gender differences in effect sizes were noted, although few studies have treated gender as an independent variable. In addition, social skills training was found to be more effective for students

exhibiting social withdrawal or learning disabilities than for aggressive students. In a 2001 meta-analysis of social skills training with antisocial and aggressive youths, Ang and Hughes found a moderate effect size for the skills training intervention. Interestingly, group composition appeared to mediate the effectiveness of the intervention. Intervention groups that were composed of deviant youth and their prosocial peers had higher effectiveness than did intervention groups composed only of deviant youth. This meta-analysis provides further evidence for the positive effect that prosocial peer models can have on social skills training interventions. Forness and Kavale's 1996 meta-analysis of 53 social skills intervention studies of students with learning disabilities found a similar positive, but less powerful, effect for social skills training. Specifically, they reported a mean effect size of only 0.21. This translates to a modest 8% percentile rank increase in performance when compared to untreated control students. Many of the studies reported by Forness and Kavale were brief interventions—less than 3 hours per week over a 10-week period.

Based on these reviews of research, there appears to be good support for the moderate effectiveness of social skills training procedures in general, and in particular for operant and modeling procedures. However, treatment effects often are short-lived and do not generalize to other settings or behaviors unless specific intervention steps are taken. For treatments delivered in social skills training to be considered truly effective and valid, levels of change must evidence treatment generality. Treatment generality is composed of maintenance (behavior change that persists over time) and generalization (behavior changes that occur under nontraining conditions). Although treatment generality is often evaluated as a "hope-for-the-best" side effect of social skills training, maintenance and generalization should be systematic facets of any intervention. Several authors have discussed procedures to enhance treatment generality. These generality facilitators include (i) targeting behaviors that will be maintained by natural reinforcement contingencies; (ii) training across different behaviors, settings, and persons; (iii) training loosely (i.e., varying stimuli such as reinforcers and tone of voice); (iv) systematically withdrawing or fading intervention procedures to approximate the natural environment; (v) reinforcing behavior change in novel and appropriate settings; (vi) including peers in training; and (vii) providing booster sessions.

5. SELECTION OF INTERVENTIONS FOR STUDENTS WITH DIFFERENT PROBLEMS

The selection of social skills interventions rests heavily on the classification of social skills difficulties as resulting from deficits in either response acquisition or response performance. Gresham and Elliott's Social Skills Rating System extended this two-way classification scheme to include areas of social skills problems, social skills strengths, and potential concurrent interfering problem behaviors. This scheme distinguishes whether a child possesses the ability to perform a target skill and whether interfering behaviors (e.g., aggression and anxiety) are present.

5.1. Social Skills Acquisition Deficits

This social skill problem characterizes children who have not acquired the necessary social skills to interact appropriately with others, or children who have not acquired a critical step in the performance of a given skill. Training in skill acquisition frequently uses direct treatment approaches, such as direct instruction, modeling, and coaching.

5.2. Social Skills Acquisition Deficits with Interfering Problem Behaviors

This social skills problem describes children with emotional (e.g., anxiety, sadness, and impulsivity) and/or overt behavioral (e.g., verbal or physical aggression, tantrums, and excessive movement) responses that interfere with skill acquisition. Thus, with social skills acquisition deficits accompanied by significant interfering behaviors, the intervention objectives are to teach and increase the frequency of prosocial behaviors while concurrently decreasing or eliminating interfering problem behaviors. Interventions designed to remediate emotional responses typically involve emotional-arousal reduction techniques, such as desensitization or flooding, paired with self-control strategies, such as self-talk, self-monitoring, and self-reinforcement. Interventions that can help reduce overt behaviors are often referred to as reductive procedures. These procedures include the use of differential reinforcement, group contingencies, and mild aversive techniques (e.g., reprimands, time out from positive reinforcement, response cost, and overcorrection). If aversive techniques are used, it is a best practice to couple aversives with socially positive procedures.

5.3. Social Skills Performance Deficits

Children with social skills performance deficits have appropriate social skills in their behavior repertoires but do not perform the behavior at acceptable levels and/or at appropriate times or in appropriate settings. Interventions for social skills performance deficits typically use manipulations of antecedent and consequent contingencies such as peer initiations, social reinforcement, and group contingencies.

5.4. Social Skills Performance Deficits with Interfering Problem Behaviors

Children with social skills performance deficits accompanied by interfering problem behaviors have acquired given social skills, but performance of the skills is hindered by emotional or overt behavior responses and by problems in antecedent or consequent control. Self-control strategies to teach inhibition of inappropriate behavior, stimulus control training to teach discrimination skills, and contingent reinforcement to increase display of appropriate social behavior are often used to ameliorate this social skills problem. Occasionally, when the interfering behaviors persist, it may be necessary to use reductive methods in addition to positive techniques.

5.5. Social Skills Strengths

Many children exhibit social skills strengths that often may be overlooked and may be undermined through lack of attention. Therefore, it is a best practice to take a proactive approach to maintaining those strengths through reinforcement of desired social behaviors. Children with strong social skills can also become part of a treatment strategy for children with social skills deficits by serving as models or as participants in peer pairing and peer initiation strategies.

See Also the Following Articles

Attention Deficit Disorders: School-Based Interventions
- Attention Deficit/Hyperactivity Disorders (ADHD)
- Emotional and Behavioral Problems, Students with
- Study Skills Instruction

Further Reading

Ang, R. P., & Hughes, J. N. (2001). Differential benefits of skills training with antisocial youth based on group composition: A meta-analytic investigation. *School Psychology Review*, *31*(2), 164–185.

Caprara, G. V., Barbaranelli, C., Pastorelli, C., Bandura, A., & Zimbardo, P. G. (2000). Prosocial foundations of children's academic achievement. *Psychological Science*, *11*(4), 302–306.

Cartledge, G., & Milburn, J. F. (1995). *Teaching social skills to children: Innovative approaches* (2nd ed.). New York: Pergamon.

Demaray, M. K., Ruffalo, S. L., Carlson, J., Olson, A. E., McManus, S., Leventhal, A., & Busse, R. T. (1995). Social skills assessment: A comparative evaluation of published rating scales. *School Psychology Review*, *24*, 648–671.

DuPaul, G. J., & Eckert, T. L. (1994). The effects of social skills curricula: Now you see them, now you don't. *School Psychology Quarterly*, *9*, 113–132.

Elliott, S. N., & Gresham, F. M. (1991). *Social skills intervention guide: Practical strategies for social skills training*. Circle Pines, MN: American Guidance Service.

Elliott, S. N., & Gresham, F. M. (1993). Social skills interventions for children. *Behavior Modification*, *17*, 287–313.

Forness, S. R., & Kavale, K. A. (1996). Treating social skills deficits in children with learning disabilities: A meta-analysis of the research. *Learning Disability Quarterly*, *19*(1), 2–13.

Gresham, F. M. (1994). Generalization of social skills: Risks of choosing form over function. *School Psychology Quarterly*, *9*, 142–145.

Gresham, F. M., & Elliott, S. N. (1989). Social skills deficits as a primary learning disability? *Journal of Learning Disabilities*, *22*, 120–124.

Gresham, F. M., & Elliott, S. N. (1990). *Social skills rating system*. Circle Pines, MN: American Guidance Service.

Gresham, F. M., & Reschly, D. J. (1988). Issues in the conceptualization, classification, and assessment of social skills in the mildly handicapped. In T. R. Kratochwill (Ed.), *Advances in school psychology* (pp. 203–247). Hillsdale, NJ: Erlbaum.

Guevremont, D. (1990). Social skills and peer relationship training. In R. A. Barkley (Ed.), *Attention deficit hyperactivity disorder: A handbook for diagnosis and treatment* (pp. 540–572). New York: Guilford.

Lentz, F. (1988). Reductive techniques. In J. C. Witt, S. N. Elliott, & F. M. Gresham (Eds.), *The handbook of behavior therapy in education* (pp. 439–468). New York: Plenum.

Mathur, S. R., & Rutherford, R. B. (1991). Peer-mediated interventions promoting social skills of children and youth with behavioral disorders. *Education and Treatment of Children*, *14*, 227–242.

Merrell, K. W. (2003). *Behavioral, social, and emotional assessment of children and adolescents*. Mahwah, NJ: Erlbaum.

Mize, J. (1995). Coaching preschool children in social skills: A cognitive–social learning curriculum. In G. Cartledge, & J. F. Milburn (Eds.), *Teaching social skills to children and youth* (3rd ed., pp. 237–261). Boston: Allyn & Bacon.

Sheridan, S. M., & Walker, D. (1999). Social skills in context: Considerations for assessment, intervention, and generalization. In C. R. Reynolds, & T. B. Gutkin (Eds.), *The handbook of school psychology* (3rd ed., pp. 686–708). New York: Wiley.

Stokes, T. F., & Osnes, P. (1989). An operant pursuit of generalization. *Behavior Therapy*, *20*, 337–355.

Social Support

Christine Schwarzer and Petra Buchwald
Heinrich Heine University, Düsseldorf, Germany

1. Introduction
2. Different Meanings of Social Support
3. Social Support and Health
4. Support Groups and Self-Help Groups
5. Social Support in the Elderly
6. Social Support and Gender
7. Social Support and Social Skills
Further Reading

GLOSSARY

emotional support Provision of liking, reassurance, and respect.
instrumental support Aid with services and tasks of daily living.
perceived support Support perceived to be available if the need for it arises in the future.
received support Support actually received if needed in the past.
social network A group of individuals with whom an individual may exchange support.

Social support is information that leads individuals to believe that they are valued, respected, and loved and that helps them to cope with major life stressors and the challenges of everyday life. Social support is a special kind of social interaction that can appear in different ways and that can be both a psychological and a tangible resource provided by a social network, for example, by friends, family members, and/or colleagues.

1. INTRODUCTION

Social support refers to the fundamental experience that people need to stay healthy both physically and mentally, to get proper aid and comfort, and to integrate morally and behaviorally into the group or society at hand. The positive effects of social contacts seem to have been well known since the beginning of humankind. Although the positive effects of social support are an integral part of a healthy life, there is much more to learn about the theoretical underpinnings, conceptualizations, and operationalizations of this construct. During the past 30 years or so of research, social support has become one of the central variables in many fields of psychology, epidemiology, public health, and education. It seems to be a potent variable for helping to explain the beneficial or detrimental effects of other people being around an individual who is coping with the challenges of daily life and/or major life stressors. But some crucial questions still remain within the community of researchers. What is the real nature of social support? When does social support occur? Are there gender differences with respect to giving or receiving social support? How can we measure social support? Is social support always helpful? What are the consequences of research outcomes in this field for applied psychologists? This article focuses on questions such as

©2004 Elsevier Inc.
All rights reserved.

these and provides an outlook for future research and practical applications.

2. DIFFERENT MEANINGS OF SOCIAL SUPPORT

Because social support occurs within the interaction of people, it is central for applied psychologists to find out what constitutes its effects. Is it just the fact of being together with others? Is it the certitude that one is loved and honored by others who console? Or is it just the helpful hands that others lend by giving financial or other tangible aid? The different viewpoints that can be found in the research literature stem from various theoretical perspectives. More than a century ago, Durkheim was one of the first to observe a higher risk of committing suicide among socially isolated individuals than among those with regular social bonds. In line with this, there exists a research tradition of looking for the objective quantitative roots of social support. This structural view is interested in the size, structure (e.g., number of males and/or females), and density of social networks and in the amount of embeddedness they give to people. However, the number of network members or the number of contacts among them is not necessarily related to psychological well-being or the subjective perceptions of being helped. Because relationships may be unhelpful as well as supportive, the focus of research has shifted from network size to an understanding of individual perception of social support as well as the context in which it is given. In contrast to the structural perspective, the more functional view is interested in relational content, which can be subdivided into perceived availability of support and actually received support. The functional view relies on self-report data of the individual either on how many helping acts (and from whom) he or she can remember for a given time in the past (received support) or on who will help in a stressful situation in the future (perceived support). Following the latter point of view, researchers are not as interested in the structure as in the content of support and focus on what people really do to each other in a network, for example, providing emotional support, informational support, instrumental support, or appraisal support. Emotional support involves the message of being loved, cared for, esteemed, and connected to other people in a network of mutual obligation. Informational support can be guidance with problem solving and advice. Instrumental support includes tangible elements such as housing, transportation, and money. Appraisal support offers personal feedback concerning personal attitudes and behavior. These various research approaches make the construct a multifaceted one that can hardly be defined in a way that incorporates all possible perspectives. But in a global sense, social support can be defined as information that people receive within social interactions, including actual assistance and feelings of embeddedness in a social network that is believed to provide love and care.

Social support can be provided by formal or informal sources. An informal and very important support source is the family (e.g., parents, spouses, other relatives). Other informal sources of support might include friends, partners, coworkers, peer groups, and neighbors. Formal social support may come from services such as self-help groups, drop-in centers for information, and formal community services. For example, schools, which often seem to be a source of stress for pupils, can also be a source of social support for children who experience problems with their parents. Hence, these formal support sources can be focused on specific populations of individuals who may be at risk, such as children and older people living in isolation, or they may be open to all people in general. Empirical studies indicate that both formal and informal support sources are positively related to health and well-being, although people seem to rely more on informal support sources for emotional help and seem to prefer formal sources for receiving instrumental support. In addition to human help, household pets can provide support. Analyses of pet support measures show that pets fulfill many of the same support functions as do humans (e.g., emotional support, social integration, closeness) and that pet support, like human support, is associated with less stress and better adjustment. However, studies that have analyzed sources of support are limited. Thus, further research should collect information about sources of social support in special groups differentiating special kinds of support.

Sometimes, social support not only is appreciated but also can be seen as a source of stress. The mere offering of support can be very stressful for a recipient if he or she does not want help from others in that particular situation. Rook was one of the first researchers to make a significant contribution to investigating the negative aspects of social support. There is some empirical evidence for the negative effects of received assumptive help, that is, help that is neither asked for nor needed. This implies a certain dilemma: If help might be appraised as threat or control, it might be better not to

offer any help. On the other hand, not offering assistance might hurt a potential recipient who really needs it. In both cases, the potential helper has reasons to worry that his or her motives will be questioned. Furthermore, if the offer of help is rejected, the helper's self-esteem may be threatened. Therefore, support providers who acquire social skills, such as knowing appropriate behaviors for certain situations, produce less negative supportive actions than do those who do not. Social skills not only improve the interactions with others but also boost the understanding of others' needs and desires. Therefore, socially skilled helpers provide support when it is needed and decrease support in situations where others prefer to help themselves.

According to the threat to self-esteem model of Nadler and Fisher, receiving help is experienced as self-supportive when recipients feel appreciated and cared for but is experienced as self-threatening when recipients feel inferior and overly dependent. If recipients feel supported, they react with positive feelings, acceptance, and gratefulness to the helpers. But if recipients feel threatened by the help they receive, this is accompanied by negative emotions and evaluations of both helpers and recipients. There are three conditions under which receiving help could be perceived as threatening. First, individuals with high self-esteem regard themselves as highly competent and are especially sensitive to the implication that they are unable to take care of themselves. As a result, they tend to react more negatively to receiving help than do those with low self-esteem. Second, being helped by other competent people highlights the contrast of being in need of assistance between helpers and receivers. Third, receiving help from significant others on important ego-relevant tasks can be threatening to recipients' self-esteem.

However, high self-esteem does not seem to prompt negative reactions to social support from those who are close to recipients. In close relationships, those negative effects of similarity do not occur because similarity is expected and desired. Feelings of inferiority are less likely to arise because mutuality in such close relationships makes receiving help less threatening. In addition, age might be an important moderator of experiencing negative effects of support in close relationships. Dependency is acceptable for children, and they react negatively to the help of their parents less often. On the other hand, older parents will not as easily accept help from their adult children.

On the basis of the previously described threat to self-esteem model, many empirical studies have been launched. Table I summarizes some outcome variables that were part of empirical studies analyzing the positive and negative impacts of social support. These studies also showed that the association between unsupportive interactions and outcome variables did not result from a confounding with the severity of an illness or the level of psychological impairment.

TABLE I
Positive and Negative Effects of Social Support

Positive effects	*Negative effects*
• High self esteem	• Low self-esteem
• Self-worth	• Inferiority
• Self-efficacy	• Helplessness
• Optimism	• Hopelessness, depression
• Joy	• Anger
• Health, life satisfaction	• Illness, psychological impairment
• Sense of belonging	• Overprotection
• Embeddedness	• Restriction of freedom, demands of conformity, social control, social conflicts
• Gratefulness	• Indebtedness to the donor
• Adaptive coping	• Dysfunctional coping

3. SOCIAL SUPPORT AND HEALTH

Since Cassel hypothesized in 1974 that, under stressful conditions, the lack of social support will lead to a higher risk of becoming ill, research has shown that supportive interactions among people are protective against the health consequences of life stressors. It appears that social support can protect people in crises from a wide variety of pathological states, from low birthweight to death. However, recent research reviews of the empirical evidence of social relationship factors being protective against life stressors have yielded rather inconsistent results. One of the main reasons for such inconsistent findings might be the rich diversity of definitions and conceptualizations of social support and the complexity of the support process. A person's ability to cope effectively depends on the stressor itself, the type of preferred coping resource, and the resources available to the individual. There are several major types of coping resources, and social support is only one of them.

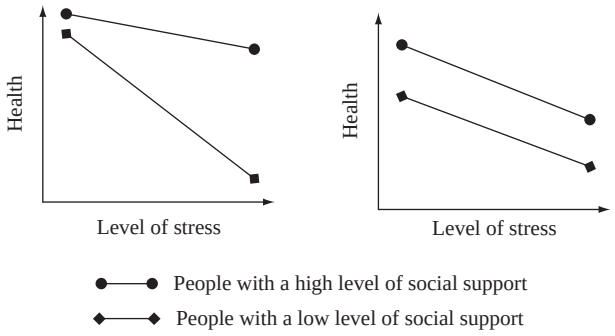

Health
Level of stress
Health
Level of stress
People with a high level of social support
People with a low level of social support

FIGURE 1 Stress buffering effect (left) and direct effect (right) of social support.

Investigating the question of stress buffering versus the direct effect provided a better understanding of the effects of social support. According to the stress buffering hypothesis, social support is needed only when people are under high stress. In that case, social support can buffer the negative stressful impact of critical life events such as divorce, loss of a loved one, chronic illness, pregnancy, job loss, and work overload on health. This means that the health and mental health benefits of social support are evident mainly during periods of high stress, whereas social support may have few physical or psychological benefits during periods of low stress. The stress buffering effect is depicted on the left-hand side of Fig. 1. Another model that explains how social support affects people suggests a direct effect and is illustrated on the right-hand side of Fig. 1. Here, social support has a direct positive effect on physical or mental health independent of stress levels. In other words, regardless of whether people are experiencing stressful situations, social support is generally beneficial.

Empirical evidence for both direct and buffering effects of social support has emerged. In 1986, Hobfoll reported that the direct and stress buffering effects occur simultaneously and that one type of effect is more prominent than the other in a given instance.

An overwhelming amount of evidence shows that social support has effects on both psychological and physical health. One remarkable result for physical health is that social support reduces the risk of dying from disease, suicide, or accidents by about the same percentage as does smoking. Moreover, increased social support from family, friends, and/or colleagues is related to decreased psychological distress. With help and support, people are psychologically better able to manage critical life events, such as unemployment and divorce, as well as everyday problems and hassles.

In Fig. 2, a theoretical framework for social support as a coping resource helps to illuminate the effects of social support on health, referring to stress theoretical underpinnings.

4. SUPPORT GROUPS AND SELF-HELP GROUPS

The rapid increase in the number of support groups and self-help groups all over the world over the past 15 years or so is likely due to rising health care costs, an increase in alternative health care, and an emphasis on prevention. Self-help groups can be seen as an attempt by people in stressful situations to enhance the amount of mutual help and cooperation they receive. They are groups of people who come together for support because they share common problems or experiences and

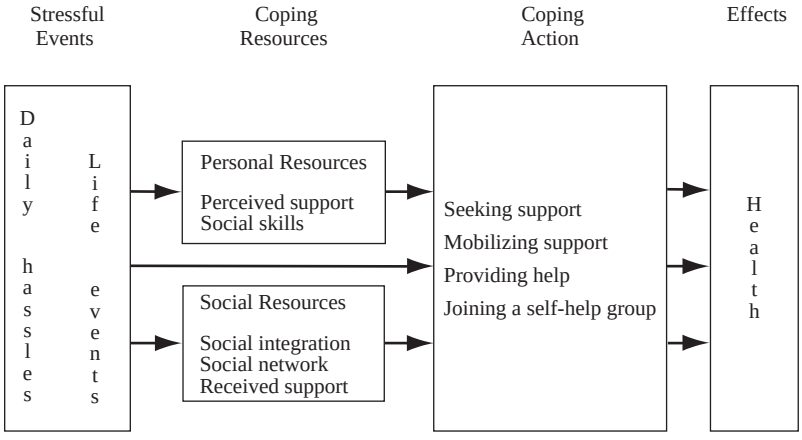


FIGURE 2 Theoretical framework for social support as a coping resource.

meet to give and receive support. Making the distinction between social support groups and self-help groups is important for practitioners in deciding which groups will best serve their clients' needs. Support groups are member centered, although they are guided by trained leaders who provide expertise in group counseling but might not share the concerns that bind the groups together. In contrast, self-help groups are mostly leaderless or guided by nonprofessionals. For example, people who are learning to deal with specific problems find support in self-help groups such as Alcoholics Anonymous. There are groups not only for alcoholics but also for former drug addicts, divorced people, single parents, cancer patients, and so on. Support groups and self-help groups both help people to cope not only because they provide assistance but also because people can learn coping techniques from others with similar problems. Recent research reveals that young unmarried White women with less social support in other areas of their lives are more likely to join self-help groups than are other individuals. Spiegel showed, for example, that women with advanced breast cancer who joined a support group not only benefited emotionally from this experience but also lived longer than did similar women who did not attend a support group.

Self-help groups can be seen as a special kind of communal coping, in contrast to individual coping efforts. Self-help groups provide advantages that are not found in individual ways of coping. During times of stress and emotional trouble, knowing that others have similar problems can be very reassuring. In addition, seeing others' problems improve can be a source of hope and motivation. Because self-help group members typically have comparable problems, they can learn from each other's mistakes and share insights. Furthermore, when a group member receives similar comments about his or her behavior from several group members, the message may be more convincing than when it comes from a single friend or therapist. Therapists often direct their patients to self-help groups to supplement their individual therapy. For example, a person with an alcohol problem can find comfort and help in a group with others who have the same problem. They exchange useful information, share their coping strategies, and gain hope by seeing others overcome or successfully manage their shared problems. Research on self-help groups for alcoholism and obesity strongly support the hypothesis of the effectiveness of self-help groups either alone or in combination with individual counseling.

5. SOCIAL SUPPORT IN THE ELDERLY

The elderly must face a life situation that is often characterized by stressful events such as the loss of loved ones, progressive health impairment, and disability that make them vulnerable to emotional disturbances. The elderly can be regarded as a risk population that is confronted with dependency, social isolation, illness, and the threat of death. Especially for people age 65 years or over who have experienced both physical and psychological limitations and impairments as side effects of aging, the beneficial effects of social support have been increasingly reported. Social support can be seen as one social key resource for the elderly, and many studies consistently show that perceived support exerts the strongest effects on health and well-being in late life. The general consensus is that those elderly persons who perceive a high amount of social support are better able to withstand the negative impact of stress than are their counterparts who receive less support. However, as the literature has shown, this process is by no means simple and social support functions in complex ways. For example, support indicators, such as the number of regular contacts and the amount of actually received social support, also play an important role in physical and emotional well-being in old age.

To enhance the understanding of how social support operates, it seems necessary to link social resources with personal resources within a resource-oriented life span model of successful aging because social resources, such as assistance, guidance, and presence provided by others, may be filtered through the existing characteristics of a person. This idea is generally confirmed by research on social support and control beliefs highlighting the social foundation of the control processes. For example, one could show that feelings of control beliefs and self-efficacy are enhanced by received social support. Perceived social support, in contrast, has weak relationships with perceived control but closer relationships with indicators of subjective health. However, received social support might be interpreted as being more central to initiating and maintaining behaviors that enable the elderly to perceive a sufficient feeling of control. Especially in old age, people's opportunities to accomplish their own performances are limited; they are handicapped, and specific abilities are reduced. This means that older people often cannot experience control in special domains, but they do have vicarious experiences, verbal persuasion, and other social influences that convey control information. By providing continued

support and enhancing control beliefs, social network members may help older people to avoid a downward spiral of resources that ultimately leads to a relapse of health. Following Baltes's theory of optimization with compensation, one can argue that older people cope with such an ongoing spiral of losses by giving up control in some life domains so as to maintain it in others. Or, they optimize control by compensation of resource loss with received support from others.

In addition, received social support is an important resource not only with regard to control beliefs but also with regard to emotions. Within one sample of senior citizens, it was observed that older people were more anxious than younger people. Low social integration and low social support were associated with higher levels of anxiety. Furthermore, it turned out that received social support, as measured by the number of visits, served as a moderator for the relationship between loss and anxiety. Those who recently experienced the loss of friends or family members and who were not regularly visited by others were the most anxious, whereas being contacted regularly served as a buffer against anxiety.

6. SOCIAL SUPPORT AND GENDER

Gender seems to make a difference as well, especially with respect to various types of support. Women are more likely than men to develop intimate friendships, to cooperate with others, and to provide support for their colleagues, friends, and loved ones. There is already broad empirical evidence for gender differences in the participation in social networks. Male participation is seemingly more extensive but less intensive. Men prefer large numbers of acquaintances, whereas their female counterparts prefer smaller numbers of intimate friendships. Differences in intensity and quantity of participation are resumed to be attributable to the activity-focused type of relationships preferred by men (e.g., sports, politics), in contrast to the emotion-focused type of relationships favored by women (e.g., chatting groups). However, such different preferences should be viewed more as general patterns than as rigid prototypes.

Furthermore, the quality and quantity of the participation in social networks seem to be a function of the life span. For example, whereas girls are principally help seekers during childhood, they develop into help providers during adolescence and adulthood.

Gender is related not only to differences in social network participation but also to differences in the willingness to seek and mobilize social support. Whereas women tend to draw on extended network support such as friends, family members, and partners, men tend to rely solely on their partners. Men's reluctance to seek social support beyond their partners might be due to the attributed femininity of emotional intimacy. Also, for relatively minor problems, men seek help less frequently than do women because mobilizing social support is more threatening to their self-esteem. Mobilization of support is also subject to life cycle and emancipation effects. Although emotional intimacy and affirmation from spouses are crucial to the emotional well-being of married men and unemployed wives, this priority is altered by female participation in the labor market. Employed women appreciate equity in the marriage in conjunction with affirmation from their spouses. However, social support entails not only the receipt of support but also the provision of support. Men and women show similar gender-specific preferences for support in that both choose more female support sources. This trend holds true not only in the private sphere but also in the public sphere, for example, in care services where women are overrepresented.

In general, gender-specific behavior concerning social support might be due to different norms for males and females in Western cultures. Part of the male sex role is to be courageous and adventurous, whereas part of the female role is to be nurturant and caring, valuing close long-term relationships. As a result, men are more likely to help in situations that demand heroic acts, and women are more likely to help in long-term relationships that involve less danger. The effects of these norms are striking. Meta-analyses have shown that men with a low level of social support are two to three times more likely to die than are men with a high level of social support. The results for women are similar but not as striking. Women with a low level of social support are one and a half to two times more likely to die than are women with a high level of social support.

7. SOCIAL SUPPORT AND SOCIAL SKILLS

People under stress have good reason to seek other people's help. The resources that the social network offers may reduce the impact of stressful events. Thus, social support is a major way in which to get resources that are not directly available for individuals. Therefore, an important field of intervention is enabling people to

receive meaningful help. What people need are social skills to communicate needs and desires and to mobilize social support when it is needed. Social skills training is widely used to train lonely people to make friends. People who have deficient social skills may learn how to act in a variety of social situations, for example, by observing others with appropriate social skills. They also might practice their new skills by role-playing with video-supported microteaching or modeling from films before applying the new behavior in real life. The results on such training have been very successful, although the main problem is to find the best skill for a special person in a particular situation with a special problem. In addition, the appropriateness will vary with gender, culture, social class, and context.

See Also the Following Articles

Elder Caregiving ■ Gender and Culture ■ Social Networks ■ Social Skills Training ■ Stress

Further Reading

Bowlby, J. (1969). *Attachment and loss, Vol. 1: Attachment*. London: Hogarth.

Buchwald, P., & Schwarzer, C. (2002). Analyzing the mutual influence of social support and control beliefs in the elderly. In C. Schwarzer, & M. Zeidner (Eds.), *Developmental issues in stress and coping* (pp. 141–163). Aachen, Germany: Shaker Verlag.

Cohen, S., Gottlieb, B., & Underwood, L. (Eds.). (2000). *Social support measurement and intervention: A guide for health and social scientists*. Oxford, UK: Oxford University Press.

Cohen, S., & Syme, S. L. (Eds.). (1985). *Social support and health*. New York: Academic Press.

Duck, S. (Ed.). (1990). *Personal relationships and social support*. London: Sage.

Gottlieb, B. H. (Ed.). (1981). *Social networks and social support*. Beverly Hills, CA: Sage.

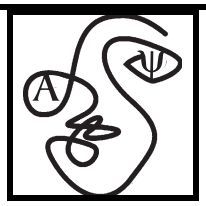
Hobfoll, S. E. (1986). *Stress, social support, and women*. Washington, DC: Hemisphere.

Sarason, B. R., Sarason, I. G., & Pierce, G. R. (Eds.). (1990). *Social support: An international view*. New York: John Wiley.

Schwarzer, C. (1992). Anxiety, stress, and received support in the elderly: A longitudinal analysis. *Anxiety Research*, *4*, 287–298.

Vanzetti, N. A., & Duck, S. (Eds.). (1995). *A lifetime of relationships*. Pacific Grove, CA: Brooks/Cole.

Vaux, A. (1988). *Social support: Theory, research, and intervention*. New York: Praeger.



Spatial Cognition

Reginald G. Golledge
University of California, Santa Barbara, California, USA

1. Introduction
2. Development of Spatial Cognition: A Multidisciplinary Effort
3. Spatial Cognition and Spatial Representation: Two Sides of the Same Coin
4. Human Spatial Cognition
5. Spatial Cognition in Animals
6. The Neurobiology of Spatial Cognition
7. Age-Related Differences in Spatial Cognition
8. Sex-Related Differences in Spatial Cognition
9. Spatial Cognition without the Benefit of Sight
10. New Frontiers in Spatial Cognition Research
Further Reading

GLOSSARY

cognition Deals with knowledge, including its acquisition, storage, retrieval, internal manipulation, and use. Stored and manipulated knowledge helps organisms achieve behavioral goals. Thus, it includes processes such as perception, learning, thinking and reasoning, memory, recall, problem solving, and language.

cognitive map A hypothetical construct implying the existence of an internal representation of an environment and the spatial relations among the environmental features.

cognitive mapping The dynamic process that involves acquiring, storing, internally manipulating, and using information about the external environment.

path integration A process by which a traveling organism self-monitors rates of travel and distance traveled as movement away from a home base takes place.

place cells The fundamental element of the hippocampal spatial mapping system; the predominant cell type in the hippocampus, comprising perhaps 90% of the cells. Place cells in the hippocampus are conceived as small, localized fields, which allow them to play a role in many spatial mappings. Place cells are part of a memory system showing location-appropriate information.

spatial cognition The internalized reflection and reconstruction of space in thought.

spatial representation An internal representation of an environment upon which an active information-seeking structure can be imposed. This internal representation of the environment enables an organism to change the location or orientation of its entire body (as opposed to perceptual, postural, and manipulatory space). Spatial representation is often used in preference to the term spatial cognition when dealing with animals other than humans.

In the multidisciplinary area of spatial cognition research, underlying theory has mostly come from cognitive and development psychology and neuropsychology. The contribution of other disciplines, such as geography, planning, architecture and design, anthropology, computer science, artificial intelligence, and information science, has often been focused on the applied facets of spatial cognition. Applications include specifying the contribution of spatial cognition to navigation and wayfinding, understanding environmental settings, building environmental schemata, learning in both familiar and unfamiliar environments, spatial knowledge acquisition across the ages, the development and measurement of

©2004 Elsevier Inc.
All rights reserved.

spatial skills and abilities, sex-based differences in spatial cognition, and cognitive maps as cultural universals. This breadth of interest enables examination of the complementary concepts of spatial cognition and spatial representations in both the human domain and the domain of other animals (particularly mammals). Questions relating to cognitive mapping and the neurobiology of spatial information processing lead to examination of age- and sex-based factors in spatial cognition and to the study of spatial cognition in disabled people. Spatial cognition is considered a necessary basis in many research contexts, and this article concludes with speculations on promising areas for research.

1. INTRODUCTION

Philosophical debates about space have continued for millennia They can be summarized into two major themes: (i) whether spatial knowledge results from experience (the empiricist position) or whether it is constructed by the mind (the rationalist or nativist position) and (ii) whether psychological spatial concepts reflect physical realities of external spaces or whether there may be striking differences between psychological and physical space.

Spatial cognition gained visibility in part because more than 100 years of experimental work in perception provided little insight about how we knew where in the world (or in the macrospace surrounding us) we were located and how we knew to select destinations and move toward them. Thus, spatial orientation and spatial representation, when added to spatial perception and spatial information processing, provided the missing links to relate human activities and knowledge to real-world objective spaces. Once established, similar processes and reasoning could be transferred to understanding the informational content of subjective (psychological) spaces. The particular contribution of different settings to environmental knowledge was acknowledged in ecological and environmental psychology as well as in urban planning, environmental design, and behavioral geography. Thus, tasks such as finding one's way around an environment, searching for objects, taking shortcuts, and recovering positional knowledge after becoming lost or disoriented all began receiving attention under this umbrella of spatial cognition.

Geographic metaphors have played a significant role in the external representation of cognized spatial information, with particular emphasis on maps, sketches, and diagrams as illustrative modes. In the early years of interest in spatial cognition, there was confusion regarding the term cognitive map. Some regarded it as a map-like way in which spatial information was stored in the brain; some used the term to describe spatial products or externalizations of stored spatial information.

In the human domain, the cognitive map system allows people to locate themselves at specific places within a frame of reference, remember what is located where, and determine how to get from one place to another in response to needs and motivations. The cognitive map also facilitates learning about environments and the spatial relations among environmental features.

Reasons for assuming that spatial cognition is a cultural universal include

- Organizational similarities of the nervous system
- Common body structures and processes, including sensory and motor activities
- Learning and socialization similarities
- A similar need to cope with complex physical environments
- Use of cognitive maps that engender spatial relations and spatial inferences
- The influence of gravity and the consequent perception of a three-dimensional world with a horizon boundary
- The influence of scale in the processing of spatial information in memory, reasoning, or language
- Classification of complex environments into categorical and hierarchical regions
- The importance of labeled features in memory and their use in problem solving
- Use of multiple frames of reference, including egocentric (body-related), exocentric (object-to-object), and environment-related
- The common understanding of angularity
- The tendency to apprehend three-dimensional spatial relations of environments by reducing them to two-dimensional representations
- The use of spatialization to produce visualizations of nonspatial information

The past two decades in particular have seen considerable research emphasis on understanding spatial cognition in humans, animals, and technical systems (computers and robots). The concept is common in many disciplines, and contributions toward understanding have come from the areas of psychology, neuroscience, philosophy, linguistics, geography, anthropology, architecture, computer science, and artificial intelligence. The result has been a rapidly growing

literature focused on activities such as wayfinding and navigation, route learning, acquisition of spatial knowledge in familiar and unfamiliar environments, strategies for regaining orientation after becoming lost, sex-based differences in spatial cognition, individual differences in spatial abilities, the development of spatial knowledge across the life span, and the building of artificial intelligences in robots that adopt strategies and perform spatial tasks either in mathematically optimal ways or in ways that closely mimic human thought and actions.

A goal common to all these areas is to understand the internal spatial representations that humans and other animals build from the experienced or imagined world. Spatial cognition goes beyond the analysis of sensory input via vision (the focus of image processing and the psychology of perception) because it includes recognition and categorization of features and objects using all the senses. Spatial concepts include, as a minimal (primitive) set, place name or identity, location, magnitude, and time. Higher order (more complex) concepts can be derived from these; the latter include spatial distribution, spatial arrangement, pattern, shape, size, texture, distance, direction, angle, reference frame, scale, link, network, region, and many others that make up a spatial concept ontology. However, spatial cognition also goes beyond object recognition to include how spatial or environmental information is encoded and how people, animals, and machines either think about space or construct representations that can be used in various reasoning processes.

Spatial cognition is revealed by performing tasks (e.g., navigation and wayfinding), by constructing external representations of stored information (e.g., spatial products such as sketch maps), by the effective use of spatial language to communicate (e.g., spoken or written spatial language), and by the ability to spatialize nonspatial information (e.g., by representing it in the form of graphs, diagrams, sketches, maps, images, art, sculpture, dance, gestures, or other modes of expression). Since it is generally accepted that vision is the spatial sense par excellence, many representations are achieved via visualizations (e.g., by diagramming). Spatial information can also be represented and used effectively in haptic and auditory domains, and a considerable portion of the natural languages used by different cultures for interpersonal communication consists of spatial concepts. These range from place names to fuzzy prepositions and prepositional phrases, such as "above," "below," "close to," "far from," "along," and "across."

Spatial information is well represented diagrammatically. Two traditional forms of representation are iconic and relational. In the iconic diagram, concrete ties among objects are established by scaled spatial relations, which are isomorphic to the originally observed relations (e.g., drawings of parts of the human anatomy). Relational diagrams are used to show spatial arrangements between entities represented by abstract, nonvisual relations, such as may be found on flowcharts, data graphs, and different topological diagrams.

2. DEVELOPMENT OF SPATIAL COGNITION: A MULTIDISCIPLINARY EFFORT

Spatial cognition is generally accepted to have emanated from Tolman's place learning theory, published in 1948. Researchers in psychology argued about differences between place learning and response learning. Other disciplines (particularly planning, design, architecture, and human geography) embraced the fundamental idea that people learned environments by encoding, storing, recalling, and using place-based information, and they took the lead in exploring ways to externally represent a person's knowledge about variously scaled environmental settings as a way to recover, represent, and use stored spatial knowledge for purposes of planning and policy making. Methods such as sketch mapping, verbal descriptions of places, and the investigation of latent spatial structures stored in memory were examined. The latter were made evident using nonmetric multidimensional scaling procedures based on paired comparison proximity or similarity data.

Spatial cognition thus developed as an interest in the way humans and other animals think about space and how they encode, store, and internally manipulate data to produce spatially relevant information that can be used in solving problems of everyday life. These problems require an understanding of spatial concepts such as location, identity, magnitude, space–time, distance, direction, orientation, frame of reference, connection, network, region, spatial pattern, shape, and distributional characteristics such as density and dispersion.

Researchers examined questions of how information was acquired and turned into knowledge. An important part of Piagetian developmental theories, as seen in the 1967 publication by Piaget and Inhelder, focused on the age-based acquisition of spatial knowledge by children. Exploring adult spatial knowledge acquisition led Siegel and White to offer a theory (called

macrogenesis) in 1975 that suggested that spatial knowledge acquisition began by noticing landmarks and then developing route knowledge as a result of experiential activity and, finally, developing relational or configurational or layout comprehension of an environment. This model influenced research on spatial cognition for more than two decades, producing theoretical and empirical models such as image theory by Kosslyn in 1975, propositional theories by Pylyshyn in 1984, dual coding theory by Pavio in 1971, anchor point theory by Golledge in 1978, and a theory of quantitative accumulation and refinement of metric knowledge by Montello in 1997. Other research, such as that by Allen in 1999, focused on an individual differences approach. Spatial knowledge acquisition thus became the building block for further developments in spatial cognition.

Macrogenesis assumes that there is development from egocentric to exocentric understanding of environments; that there is a transition in the geometries needed to comprehend space from topology to projective to Euclidean forms; and that landmark knowledge dominates in the early phases, followed by route-based knowledge with survey or configurational knowledge emerging as the ultimate abstract and complex knowledge structure. Critics of this macrogenetic model have argued that it fails to take account of knowledge acquisition and loss across the life span as aging occurs, that Piaget's stage of development theory fails to adequately consider individual differences, and that it is not an appropriate model on which to base explanations of adult learning of new environments (because adults already have the ability to think metrically and abstractly about spatial relations). Moreover, there has been constant discussion about the relative importance of landmark vs route-based knowledge acquisition. Some argue that landmark recognition must occur first for locations, anchor routes, and configurations, whereas others argue that most of our knowledge is gained from experiencing space via movement and, therefore, route knowledge must necessarily be the initial and primary source of spatial knowledge acquisition. No consensus exists.

All points of view, however, acknowledge the importance of a frame of reference in the process of acquiring spatial information. It is generally accepted that there are two dominant modes: egocentric, in which environmental information is related to the location of the knowledge seeker, and exocentric, consisting of the spatial relations embedded in the layout of objects regardless of the presence or location of a knowledge seeker.

Attaching nonspatial information to some type of locational or, more generally, georeferenced place is an example of spatialization. This is the process of making nonspatial things spatial for ease of representation, interpretation, and use (e.g., plotting age and income data on a graph).

3. SPATIAL COGNITION AND SPATIAL REPRESENTATION: TWO SIDES OF THE SAME COIN

A significant focus within the realm of spatial cognition is the process of cognitive mapping. Evidence of the presence of a cognitive map includes the ability to perform wayfinding activities by following new routes that can include detours and shortcuts. The existence of a cognitive map implies that wayfinding can take place by piloting or by the homing vector activity defined in path integration. Thus, given a current location, a new route can be deduced from memorized spatial relationships between the current location and other locations (egocentric organization) or between different locations (exocentric processing, often associated with imagination conditions).

Research indicates that spatial knowledge only reaches the required degree of coherence to be called a cognitive map in primates. The less loaded term spatial representation is often used for species other than primates. In the latter case, spatial representation is often described in terms of locomotor space, in which an organism evinces an ability to change the location or orientation of its entire body with respect to the surrounding environment. The concept of spatial representation has been tested extensively by examining location behavior with respect to biological needs. Locomotion behavior relies on developing correspondence between information stored in spatial memory and information perceived and encoded about the current task situation.

A cognitive map consists of declarative knowledge. This knowledge is organized and manipulated by cognitive processes during thinking and reasoning. Bits of information relative to some decision-making, choice, or problem-solving situation are manipulated via procedural rules, propositional structures, image schemata, and logic and other modes of thinking and reasoning.

Cognitive maps or spatial representations have several functions. First, they serve as a reference for

comparing current inputs with data stored in memory to detect matching or mismatching of schemata (such as might occur in changed environmental settings). Second, they contribute to the formation of structures or abstract representations that may involve general spatial properties common to a variety of situations. Finally, they facilitate or enable computation and planning of oriented trajectories, such as wayfinding and search within the external physical world. Thus, spatial representations may be regarded as occupying the interface between memory and action.

4. HUMAN SPATIAL COGNITION

Humans are born into a three-dimensional spatial world. In adapting to and living within this world, three spatial reference frames are used: body-centered (egocentric), object-centered (exocentric), and environment-centered (landmark-based). For humans, who constantly perceive the spatial environment as a layout of landscape features, it is possible that our general representation of locomotor space owes as much to our apprehension of space through experience, internal feedback, and optic flow during locomotion as it does to the direct perception of spatial objects scattered throughout the environment. To determine location at any particular moment, spatial updating is essential. This takes place with respect to objects in the environment and with respect to a related frame of reference.

The following are basic assumptions of human spatial thinking:

- Geographic knowledge representations in the human mind are constructed on demand.
- A representation is constructed based on knowledge stored in long-term memory.
- Information in long-term memory is fragmentary and hierarchically structured.
- A representation constructed in working memory can be visual, propositional, or spatialized.

Egocentric relations are related to a body framework, such that front–back and left–right axes are important, as is the up–down axis involved in gravitational force on the body. Exocentric frames are based on object-to-object relations that are often specified in terms of arbitrary coordinate systems, interpoint distances and directions, layout patterns or configurational structure, and layout or object-by-object relations to a superimposed, external frame of reference (e.g., a street system or the latitude and longitude coordinate system). As a body experiences an environment, sensory input is included with respect to one or another of these frames of reference. There is ongoing discussion as to whether spatial information is orientation free or orientation dependent, whether (or to what extent) alignment forces influence encoding processes, and the extent to which information is view (or perspective) dependent. One pervasive question concerns how information coded in separate environments is integrated by the mind to form spatial relational structures (e.g., networks from individual routes). There appears to be strong evidence that spatial information is stored in memory in a hierarchical fashion that facilitates recall, decoding, and use. There is also evidence of a regionalization effect for spatial information, in that distance estimates between pairs of points are judged smaller for distances within a region than similar interpoint distances between places in different regions. Smaller distances tend to be overestimated and longer distances underestimated (regression toward the mean), and interpoint distances closer to a viewpoint are overestimated relative to interpoint distances far from a viewpoint. Research findings include that distances toward landmarks are often underestimated and distances from landmarks are overestimated, and that distances to landmarks are underestimated in comparison to distances to nonlandmark features. This throws doubt on the validity of a metric assumption (i.e., that distances between points are symmetric) in the world of spatial cognition.

To help us think about and communicate about spatial information, we use two kinds of tools to assist us: visualizations of phenomena (that can be represented on maps and pictures and so on) and spatializations of information (that are more metaphorically visual, including graphics, charts, and diagrams).

5. SPATIAL COGNITION IN ANIMALS

Exploratory behaviors in mammals are reactions to novel settings. This is easily observable in a variety of animals, but most experimental work in psychology, for example, has focused on rodents or primates. Both engage in wayfinding activities and reveal specific patterns of movement-related behaviors. In rodents, exploration usually follows a short freezing phase in which the various senses are used to intake available information and evaluate and look for danger. Exploration follows, and multiple contacts with objects

located in the environment take place via the phenomenon of habituation. This is part of a more general process of knowledge acquisition. In settings previously encountered, reexploration or dishabituation occurs. Thus, novelty in an environment exists normally through reference to stored memories of previously explored or familiar places. The key here is that if novelty is detected and exploration behaviors are triggered, this implies that some representation of the initial situation has been constructed. Such exploratory behavior is assumed by neuropsychologists to be an updating activity of the currently stored cognitive map or spatial representation. When experiencing a new environment, a mismatch between what is stored in the cognitive map or spatial representation and what is observed triggers exploration. In this context, spatial cognitive abilities in animals are often evaluated in a reaction-to-change environment, when repeated exposures to a given spatial configuration are followed by habituation or exploration to fix spatial relationships between environmental objects. If an object is displaced, then the spatial arrangement changes. New exploratory activity provides evidence of a mismatch between the new representation or arrangement and a representation (or cognitive map) of the original situation. Spatial processing and learning in animals are often examined via shortcut and detour tasks (e.g., when a learned path becomes blocked and requires some alternate route to be selected to reach a goal).

Much animal wayfinding activity is said to be governed by the process of path integration. This type of activity has often been observed in the behavior of ants, bees, wasps, or other nesting or home-based organisms. Thus, when foraging behavior occurs and food is found, the organism can return in a direct line or shortcut to the home base. Often called dead reckoning, this action is taken to mean that spatial representations occur across species from insects and other invertebrates to mammals. In insects, the spatial representation may be produced as a retinoptic image (or snapshot) compared to the constructed representations of mammals.

Homing has been observed in a wide variety of species, including ants, spiders, dogs, bees, wasps, and humans, suggesting that it may involve a similar algorithm for processing route-based information across these different species. This appears to be true even when the input data vary among the different tasks. For example, day roving species may use visually detected landmarks to correct minor errors in homing. However, when visual cues are not directly available (e.g., in night active rodents), when deprived of any directional reference, or when animals are prevented from using the sun's magnetic field (as when a reflective surface is interposed), they still perform homing tasks successfully, implying that they assess distance and direction through inertial signals from the vestibular system, somatosensory feedback, and motor signals that specify speed of locomotion and change of direction. Vestibular signals are presumed to play a key role in the estimation of rotations but not translations. For day foraging rodents, it has been suggested that dead reckoning is mainly a backup system that allows an animal to explore novel or unfamiliar parts of an environment when access to familiar landmarks is not available.

6. THE NEUROBIOLOGY OF SPATIAL COGNITION

The study of spatial cognition in cognitive neuroscience focuses on identifying the neural and computational mechanisms associated with spatial activities such as path integration, wayfinding, exploring imaginary environments, or locating invisible targets from memory. Cognitive neuroscientists have investigated the possible existence of place cells for providing location-based object recognition, language production and understanding of location- and direction-based verbal statements, navigation, and orientations in space. Cognitive psychologists suggest that the occipital lobe, the parietal lobes, the dorsal system in general, and the hippocampus are essential to the encoding, storage, processing, and decoding of spatial information.

It has been proposed that spatial information is stored in the hippocampus and neocortex, where spatial learning and memory take place. This hypothesis was popularized by O'Keefe and Nadel in 1978. The hypothesis results from examining the time course of the respective involvement of the hippocampus and the neocortal areas in learning and memory tasks. The argument proposes that spatial memories are initially stored in the hippocampus and become gradually incorporated into the neocortal system as structures or abstract forms. To prevent interference with the previously stored memories in the neocortical system, knowledge structures need to build up slowly, along with experience. The term structure is essentially equated to the term cognitive map. Thus, cognitive mapping can be readily seen to be a process for orienting and for active information-organizing activities in the brain.

The seminal work by O'Keefe and Nadel helped formulate the notion of place cells as the fundamental elements of a hippocampal spatial mapping system. The importance of the hippocampus in spatial thinking and reasoning has been borne out by research on rats with the hippocampus removed and by some research on brain-damaged humans. Given a task that can only be solved using a place strategy, consequent hippocampal dysfunction makes spatial learning impossible. Other research has produced evidence that animals without a functional hippocampus solve spatial tasks predominantly with response or guidance strategies rather than place strategies. Recent research appears to have borne out these predictions, particularly research involving the use of the Morris water maze.

7. AGE-RELATED DIFFERENCES IN SPATIAL COGNITION

The study of the development of spatial knowledge structures in children, known as developmental theory, has a long tradition of Piagetian influence. Much less research has been undertaken on spatial cognition in the aged, which has been limited to experiments examining the effect of declining information processing speed on visuospatial abilities. However, with the increasing aging population of the post-World War II baby boomers and with more attention being paid to memory loss (Alzheimer's disease), the behavior of the elderly and concerns about the general quality of life of aged people have become a more popular and pressing applied area for spatial cognition research.

The aged population (generally defined as those people 65 years of age or older) comprises approximately 15% of the total U.S. population and will grow to approximately 20% in the next few decades. A high proportion of this population lives independently, navigates, and drives to work, to go shopping, and for leisure, health, or other needs. There is legitimate concern about gaining information relating to continued spatial competence. In addition, a very high proportion of the aged population has severe vision problems (approximately 80% of the vision-impaired population in the United States is aged 65 or older). This again emphasizes the need to know the state of age-based spatial competencies (skills and abilities) required for safe and effective life in the spatial domain. Research needs to examine questions of how elderly people interact with spatial situations and how individual differences mitigate the circumstances of such interactions. This requires knowledge about information processing abilities of the aged, the variety and extent of physical abilities present in this group, neurological states, memory loss, and information retrieval and response times. For example, in learning about places, older adults tend to recall significantly fewer landmarks than younger adults, and their memory of most routes (measured in terms of retrace abilities or route sketching abilities) is generally less than that of younger populations. One interesting research finding is that there is an increased importance of landmarks in wayfinding and in spatial cognition generally among the aged. Apparently, elderly people perform wayfinding tasks more effectively if they have prior information about the area to be explored (e.g., by developing a travel plan). However, in general, research shows that route learning ability declines with age. This means that route retracing, landmark recognition, and sequential ordering of landmarks are tasks that are performed more poorly by the elderly than by younger participants.

The differences between aged and younger people are most marked when dealing with unfamiliar environments. In familiar environments, many spatial information-processing tasks are performed quite well by the elderly, and the difference between the aged and the young diminishes significantly. Evidence of spatial cognitive abilities (e.g., distance estimations and judgments of spatial extent) shows systematic age-related distortions consistent with the effort that is perceived to be required to interact physically with the test area.

8. SEX-RELATED DIFFERENCES IN SPATIAL COGNITION

Evolutionary psychologists and anthropologists agree that from the earliest times both sexes have been required to be spatially aware and able to act with reflex speed in certain (life-preserving) spatial settings. Sex-based differences in spatial abilities may have evolved over time as a function of a historical division of labor. This "hunting" versus "food gathering and home care" dichotomy is said to be reflected in substantial differences in the diameter of male and female home ranges. Because of the many dangerous situations in which they were involved as hunters, it is argued that males developed an ability to integrate novel incoming spatial information quickly and tended to develop reference frames that extended beyond the visual domain. Females are

assumed to have developed more localized spatial skills associated with foraging for food. These skills involve recognition and recall of landmarks, noticing and storing the details of local environmental settings, understanding configural relations among local environmental features, and developing incidental memory of the location of objects. It is also argued that females are able to better record details of their immediate environment through an unconscious imaging process.

There have been significant differences in the form of testing for sex-based differences in males and females. Some testing has relied exclusively on psychometric spatial ability tests (e.g., the Water Level Test, Primary Mental Abilities Test, Shepard–Metzler Rotation Test, Vandenberg–Kuse Test, and Silverman and Eals Test), which have produced results that appear to indicate that males almost universally outperform females on three-dimensional rotation tests and sometimes on other spatial visualization or surface deployment tests. Contrary results (except for the rotation tests) are often found in the literature, and reports indicate that on some spatial tests (e.g., Silverman and Eals), women tend to outperform males.

It is traditionally believed that spatial abilities are higher in men than in women. As societies and cultures have evolved, as equal opportunity in all facets of social, economic, and recreational life has become more the norm, and as unisex participation in spatially related activities (e.g., sports) has occurred, the gap between male and female scores on spatial ability tests has narrowed. It is now believed that more effacing self-stereotyping by females has played a role in influencing performance on spatial tasks (whether laboratory or field oriented).

Wayfinding research has indicated no significant sex-based differences on tasks such as accuracy of pointing to landmarks, on wayfinding by visitors to an office building, and on wayfinding by the blind. However, males have been found to be more accurate than women with regard to the geometric (coordinate) placement of objects on a map, defining the direction of landmarks or other objects in environmental settings, estimating travel distance, and using cardinal reference points to give directions. Researchers have also found that males make fewer errors and require fewer trials to learn a novel route and perform better (fewer errors) in computer simulations of wayfinding. On the other hand, women have been found to more likely refer to landmarks when giving travel directions and are more accurate in the recall of specific landmarks or objects referenced to landmarks. It has been suggested that males more readily keep track of current location with respect to distant reference points (implying the use of configurational understanding), whereas women prefer landmark-based, step-by-step instructions (often called piloting) to direct them along a route or to recall it. Studies of the use of technology have indicated a trend toward masculine dominance in the use of wayfinding technologies and that spatially based video games are less popular with females than males. When maze-based computer games are emphasized, males solve the maze problems significantly faster than do females. In general, males outperform females on the Vandenberg Mental Rotation Test, the Shepherd–Metzler Rotation Test, the Guilford–Zimmerman Spatial Orientation Test, and the Money Roadmap Test. However, there is evidence that male/female differences are observed more when encoding involves only a single task; where simultaneous encoding of dual tasks is required, sex-based differences are more difficult to observe.

Females apparently notice more details in a given environmental setting than do males. Results from the Silverman and Eals Test support this finding, as do results from other tests that involve recalling the number and location of multiple objects on a computer screen. This applies to both intentional and incidental learning scenarios. Males tend to outperform females on tests of general geographic knowledge. This is supported by evidence that many more males than females make it to the finals of the National Geography Bee, and that male scores on geographically relevant questions on college testing exams are higher on average than female scores.

9. SPATIAL COGNITION WITHOUT THE BENEFIT OF SIGHT

Without vision, a person has no visual memories of particular places and has to substitute other senses for direct visual input while experiencing environments. Also, such persons must cope with the fact that fundamental distance, direction, orientation, and reference frame information that is almost automatically provided by vision is absent and must be substituted for. At times, spatial experience comes from hearing; at other times, it comes from touch and motion. Vision provides relatively simultaneous perception of the macrospace—simultaneous with experience. Any objects, their locational positions, and the spatial relations among them can be absorbed in a few glances.

Even when attention wanders, sight continues to operate and incidental information is noticed and encoded. Important cues such as size, dominance of form, color, or specific function are easy to discern with sight and very difficult to discern without it. Sight allows a precise geometry to be used to help define spatial relations, and sight can provide access to information about far more distant places than can be obtained from the other senses.

In small-scale (local body) space, haptic exploration using the hands and arms provides a way to explore object location. The body to which hands and arms are connected provides a stable egocentric frame of reference, even when moved around a tabletop setting, for example. There is evidence that those who have lost their sight later in life and have different periods of sighted experience can more easily comprehend spatial concepts and, consequently, their performance on a variety of spatial tasks is quite similar to those of sighted people.

Evidence of spatial directional knowledge is usually obtained by pointing. Pointing can take place with a wand, a finger, or an arm, or the entire body can be rotated to point in the direction of an object. This process does not require sight and is often used in studies of spatial updating. Spatial updating in the sighted and the adventitiously (late) blind group has been shown to be reasonably accurate in locomotion conditions but less accurate in imagination conditions. In contrast, early (congenitally) blind subjects seem to perform similarly in both conditions.

10. NEW FRONTIERS IN SPATIAL COGNITION RESEARCH

Applied areas in which spatial cognition is of increasing importance include wayfinding and navigation, the development of Geographic Information Systems, remote control of robots, teleoperation, location-based services, in-vehicle navigation, and building smart environments—particularly using visualization techniques for presenting relevant spatial information in a decision support system context.

Recent research in spatial cognition has emphasized the impact of alignment effects on spatial recognition and information use, suggesting a major distinction between egocentric heading or "track up" perspective in map displays for in-vehicle navigation systems and the exocentric display of such information in orientation-fixed ("north up") displays. In general, research suggests that when there is a matching perceptual alignment between the intrinsic geometrical structure of an environment and the orientation of a representation, it leads to more accurate, more easily accessible mental representation. Such results are impacting how assistive navigation technology is being installed in private and public vehicles for transit and transportation purposes.

The need for communication between people about places became more pervasive as the technological revolution emerged and accelerated. It produced the need to discover the nature of spatial information as communicated in spatial language, to comprehend the value of the processes of spatialization (e.g., graphing or mapping) or making representations (e.g., photos or images), and to understand layouts (such as can be captured by all or any of these representational processes). Interest has been increasing at an exponential rate.

In the future, it may be feasible to conduct more extensive research on relationships between aging and spatial cognition using virtual environments and magnetic resonance measuring techniques for examining the effect of brain cell loss on human spatial performance. Obviously, as the aged group increases in size, the critical question of accessibility in the general spatial–geographic domain will become even more important. Examining the extent of spatial experiences of the aged should be a focus of socially relevant applied psychology in the future.

Spatial cognition has provided underlying algorithms for the incremental generalization of routes that are independently traveled by a mobile robot, which then can be integrated into a spatial representation (or network) of these routes. In robot-related artificial intelligence, a route is defined as a sequence of decision points that are connected by line segments along which a robot can travel. Robot navigation consists of a series of spatial language instructions that, when followed in sequence, allow the robot to select links, make turns at specific angles, and proceed to a destination.

Spatial cognition occupies a place of growing importance in the general area of location-based services and mobile computer applications. This includes design and use of next-generation smart phones, personal digital assistants, wearable computers, and modifiable portable computers such as Palm Pilots. Design of these technologies in such a way as to facilitate use by naive observers has meant an increasing emphasis on the use of concepts and theories related to spatial cognition. In particular, one needs to know what cognitive processes are involved in using them and how to

minimize this cognitive effort to facilitate naive usage. Concerns range from how to present location information to a potential user to how to represent environmental layout information in an interpretable way on the small LCD screens typically built into such portable devices. For example, if maps are to be a primary communicative device, then questions of scale, simplification, generalizability, alignment, and symbology will become important issues.

See Also the Following Articles

Cognition and Culture ■ Cognitive Aging ■ Motives and Goals

Further Reading

Allen, G. L. (1999). Spatial abilities, cognitive maps, and wayfinding: Bases for individual differences in spatial cognition and behavior. In R. G. Golledge (Ed.), *Wayfinding behavior: Cognitive mapping and other spatial processes* (pp. 46–80). Baltimore, MD: Johns Hopkins University Press.

Burgess, N., Jeffery, K. J., & O'Keefe, J. (Eds.). (1999). *The hippocampal and parietal foundations of spatial cognition*. New York: Oxford University Press.

Devlin, A. S. (2001). *Mind and maze: Spatial cognition and environmental behavior*. Westport, CT: Praeger.

Freksa, C., Brauer, W., Habel, C., & Wender, K. F. (Eds.). (2000). *Spatial cognition II: Integrating abstract theories, empirical studies, formal methods, and practical applications*. Berlin; New York: Springer.

Gallistel, C. R. (1990). *The organization of learning*. Cambridge, MA: MIT Press.

Gärling, T., & Golledge, R. G. (Eds.). (1993). *Behavior and environment: Psychological and geographical approaches*. Amsterdam: Elsevier/North Holland.

Golledge, R. G. (1978). Learning about urban environments. In T. Carlstein, D. Parkes, & N. Thrift (Eds.), *Timing space and spacing time, Volume I: Making sense of time* (pp. 76–98). London: Arnold.

Golledge, R. G. (Ed.). (1999). *Wayfinding behavior: Cognitive mapping and other spatial processes*. Baltimore, MD: Johns Hopkins University Press.

Kimura, D. (1999). *Sex and cognition*. Cambridge, MA: MIT Press.

Kitchin, R., & Freundschuh, S. (Eds.). (2000). *Cognitive mapping: Past, present and future*. London: Routledge.

Lloyd, R. E. (1997). *Spatial cognition: Geographic environments*. Dordrecht, The Netherlands: Kluwer.

Montello, D. R. (1997). The perception and cognition of environmental knowledge: Direct sources of information. In S. C. Hirtle, & A. U. Frank (Eds.), *Spatial information theory: A theoretical basis for GIS* (pp. 273–311). Berlin: Springer-Verlag.

Morris, R. G. (1981). Spatial localization does not require the presence of local cues. *Learning and Motivation*, *12*, 239–260.

O'Keefe, J., & Nadel, L. (1978). *The hippocampus as a cognitive map*. Oxford, UK: Clarendon.

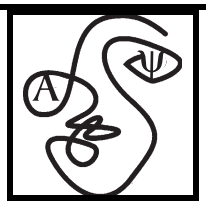
Passini, R. (1992). *Wayfinding in architecture*. New York: Van Nostrand Reinhold.

Portugali, J. (Ed.). (1996). *The construction of cognitive maps*. Dordrecht, The Netherlands: Kluwer.

Pylyshyn, Z. W. (1984). *Computation and cognition: Toward a foundation for cognitive science*. Cambridge, MA: MIT Press.

Talmy, L. (2000). *Towards a cognitive semantics* (Vol. 1). Boston: MIT Press.

Talmy, L. (2003). *Towards a cognitive semantics* (Vol. 2). Boston: MIT Press.



Special Education

Stephen D. Truscott, Ruth A. W. Berry, and Kelvin Lee

University at Buffalo, State University of New York, Buffalo, New York, USA

1. Introduction
2. History and Theoretical Foundations: From Institutionalization to Inclusion
3. Current Special Education Practices
4. Medical and Technological Progress
5. Multicultural and Racial Issues in Special Education
6. International Perspectives
Further Reading

GLOSSARY

continuum of services Special education services ranging from limited support (e.g., physical therapy, counseling) to full-time, out-of-state residential programs.

educational disabilities Any of the 13 disability categories identified by U.S. law as eligible for special education (e.g., learning disability, mental retardation, emotionally disturbed, deaf, other health impaired).

free and appropriate public education (FAPE) Education provided as a federal civil right to all students between 3 and 21 years of age regardless of disability; education must be free to the students, must meet identified needs (i.e., potentially beneficial) appropriately, and must be as near to general education as possible.

general education Public education provided to all students.

inclusion Special education services provided in the general education setting; philosophically, inclusion exceeds least restrictive environment (LRE) requirements.

individualized educational program (IEP) Program designed for all students in special education; it (a) documents current levels of performance, learning, social, physical, and behavioral needs; (b) outlines a program to meet these needs; (c) identifies any special services, testing modifications, and/or equipment needs; and (d) provides annual goals and objectives.

least restrictive environment (LRE) Requirement that special education must be provided to the maximum extent possible in or near the general education placement.

special education Supplemental or replacement educational services designed to address the educational, social, physical, behavioral, and preoccupational needs of students with recognized educational disabilities.

Special education is a service provided to students with educational disabilities. This article provides a general overview of the history, current practices, and contemporary issues regarding special education. It focuses on special education practices in the United States, although other developed nations implement comparable practices and share similar concerns. It also provides a brief review of international issues in special education.

1. INTRODUCTION

In 1896, Lightner Witmer opened one of the first psychological clinics in the United States. His first client was a child referred for spelling difficulties by a public school teacher. In 1907, Witmer wrote that the clinic served children who were "conspicuous because of an inability to progress in school work ... or [who had] ... moral defects which rendered them difficult to manage." The

©2004 Elsevier Inc.
All rights reserved.

partnership between applied psychology and special education began early and continues to this day. Over the years, psychologists have often led efforts to find effective educational, behavioral, and psychoeducational interventions to address the needs of students with disabilities. Most recently, 5 of the 19 members of the 2002 President's Commission on Excellence in Special Education (PCESE) were psychologists. This article provides a general overview of the history, current practices, and contemporary issues regarding special education and its relationship with applied psychology.

2. HISTORY AND THEORETICAL FOUNDATIONS: FROM INSTITUTIONALIZATION TO INCLUSION

People with disabilities have always been part of society. What has changed is how society defines and explains these differences and how people with special needs are treated. Early explanations of disabilities were grounded in superstitious belief systems, and treatments included abandonment and extermination. By the 1800s, these gave way to quasi-scientific explanations and institutionalization of people with disabilities. Early efforts to educate children with disabilities were often conducted by physicians (e.g., Itard, Seguin, Howe, Gallaudet) whose work established the foundation for current practices such as individualized instruction and an emphasis on functional skills. Because societal demands were different (e.g., for literacy), early efforts to provide special education were focused on people with severe disabilities such as mental retardation and psychosis. During this period, asylums for the treatment of individuals with disabilities began to appear. Such institutions were prevalent in the United States until the relatively recent past and still exist widely in many other countries. Although asylums were originally intended to protect people with disabilities from cruel treatment in the outside world, patients were sometimes abused and neglected.

2.1. The Child Study Clinic Movement

Witmer's psychological clinic was part of a movement that featured the application of scientific methods to problems of living, an emphasis on public health, and attention to social justice. At the turn of the 20th century, such "child study clinics," often associated with public schools, sprouted up in metropolitan areas throughout the United States. These clinics focused on using the new science of psychology to address the problems of children. The clinics served students with a range of disabilities, including mental retardation, learning problems, and behavior disorders. In addition to prescribing treatments and educational interventions, these pioneering applied psychologists viewed schools as places to provide things such as nutrition, parent training, and healthy environments when these were missing at home. It is no coincidence that child study clinics were established at approximately the same time as the rapid expansion of public education and the implementation of child labor laws.

The clinics focused on applying the new science of psychology to the needs of individual children, so it is not surprising that the Binet–Simon test (translated in 1910) found widespread acceptance among early clinicians. The new test provided an objective and scientific way in which to classify students and played a significant role in the rise of special programs for students with disabilities in metropolitan areas served by clinics. During the 1920s, special programs for students with disabilities (notably mental retardation) were widespread in cities but nearly absent in rural America. The Binet test, administered by clinicians in child study clinics, played a critical role in these placement decisions.

2.2. The Rise of the Medical Model and the "Within Child" Deficit Approach to Special Education

Throughout the 20th century, psychological and medical technologies were increasingly evident in special education diagnoses and programming. "Within child" and biological explanations of disabilities (e.g., minimal brain dysfunction, dyslexia, Down syndrome) became more and more dominant after 1950. Following the medical model, disabilities were thought to originate within the child and were most often attributed to manifestations of underlying biological problems. Special education treatment followed along with these explanations, and the idea of diagnostic prescriptive teaching came to the forefront. This model is predicated on the idea that particular diagnoses can be addressed best by prescribing specific treatments. In special education, the individualized educational program (IEP) exemplifies this approach.

Concurrently, after the 1954 *Brown v. Board of Education* U.S. Supreme Court decision ruled against

public school segregation by race, a series of court cases and legislation provided U.S. constitutional rights (e.g., equal protection) to people with disabilities. Special education rights grew rapidly, culminating in 1975 with the enactment of Public Law (PL) 94-142, a federal law that granted access to a free and appropriate public education (FAPE) to all students regardless of disability. The rapid expansion of special education services after PL 94-142 has continued to date.

2.3. Current Conceptualization: Behavioral, Ecological, and Sociocultural Models

Currently, behavioral, sociocultural, and ecological approaches mark a fundamental transition in special education in the United States. In contrast to the medical model, these contemporary approaches attribute the manifestation of a "disability" to the transaction between the demands of the environment and the behavior of the individual rather than to the expressions of internal biological deficits. Fueling the transition is the understanding that the widespread goal of recognition and inclusion of people with disabilities in the educational system has been reached but that less attention has been focused on the learning, social, behavioral, and occupational outcomes of special education placement. In addition, considerable growth has occurred in the percentage of students classified in the mild or "judgmental" categories of disability (i.e., those disabilities that are determined primarily by results of psychometric testing such as learning disability [LD], speech/language impaired, and emotionally disturbed). Such students now comprise 85 to 90% of the special education population in the United States. The percentage of students in disability categories with a known etiology or physical problem has remained much more stable over time.

Indeed, in the United States today, most students in special education have neither clearly defined disabilities nor known etiologies. Wide differences among the characteristics of students within the same disability classification in various settings support the idea that most educational disabilities are primarily social constructions. For example, in 1994, Gottlieb and colleagues reported that the average intelligence quotient (IQ) of students labeled as LD in urban schools was 81, compared with an average of 102 for students labeled as LD in suburban communities. Examination of such differences has resulted in the serious ongoing reexamination of special education definitions and services exemplified by the PCESE's finding in 2002 that many students in special education are essentially "instructional casualties" rather than students with educational disabilities. As programs based on the medical model (e.g., neurological models of LD) have failed to result in acceptable outcomes for the majority of students in special education, new interest in functional assessment of behavior, direct assessment of academic skills (e.g., curriculum-based measurement), attention to observable intervention outcomes, and direct instruction of academic, social, and behavioral skills has become an important element of special education reform in the United States. These practices, which are derived from behavioral psychology and ecological developmental theory, are based on the idea that the learning problems experienced by most children labeled as educationally disabled arise from a mismatch between the learning needs of the students and the instructional or other environments.

3. CURRENT SPECIAL EDUCATION PRACTICES

3.1. Classification and Labeling

In the United States, approximately 13% of public school students were enrolled in special education programs during the 2001–2002 school year. The largest disability group, nearly half of the total, is composed of students identified as having specific learning disabilities (SLDs) (Fig. 1). The next largest subgroup is composed of students with speech or language impairments. Although the vast majority of students in special education have mild difficulties, educational disabilities can range in severity from mild to profound and can range in scope from narrow to broadly pervasive. For example, a learning disability may affect only a student's ability to complete mathematical calculations, whereas a child labeled as multiply disabled may have severe cognitive, physical, social, and behavioral difficulties. Both of these students may require special education, yet the scope of the disability, required interventions, and educational goals for these two children would be quite different. Educational disabilities are caused by a range of factors that can be located on a continuum from environmental deficiencies (e.g., lack of effective reading instruction) to within child physical factors (e.g., genetic defects). In general, the more severe the disability, the more likely it is that the etiology can be identified.

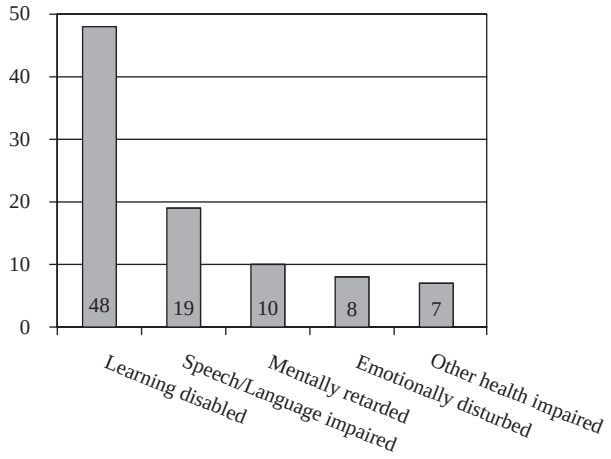


FIGURE 1 Approximate percentages of U.S. special education students in the 5 major disability categories (2002 data). Approximately 92% of all U.S. students in special education are classified in these 5 of the 13 educational disability categories.

In the United States, definitions of educational disabilities are specific to special education and are separate from "disabilities" or diagnoses used in other legislation (e.g., Americans with Disabilities Act) or systems (e.g., *Diagnostic and Statistical Manual of Mental Disorders*). Special education definitions require two primary components. First, it must be determined that a student has a need for special education. That is, the suspected disability must negatively affect the child's education. Second, the suspected disability must be in one of the following general categories: (a) sensory disabilities such as visual impairments, hearing impairments, and deafness–blindness; (b) physical and neurological disabilities such as orthopedic impairments, other health impairments, traumatic brain injury, multiple disabilities, and autism; and (c) developmental disabilities such as SLD, speech or language impairments, emotional disturbance, mental retardation, and developmental delay. Children up to 9 years of age can be identified as having a developmental delay, reserving the identification of specific disabilities until they are older and the disability characteristics become more distinct.

3.2. Identification Processes

In most U.S. schools, determining a student's need for special education begins with a teacher referral and prereferral intervention. Prereferral intervention most often involves a team of administrators, specialists, and teachers that examines the student's educational difficulties and recommends interventions or services to address the student's needs in the general education setting prior to special education referral. Ideally, such a team implements a problem-solving approach by collecting relevant data, recommending targeted general education interventions or services, and monitoring the results of these recommendations. Over time, if the team concludes that the child has not benefited sufficiently from the recommended interventions, the child is referred formally for a special education evaluation.

The formal determination of an educational disability is made by a multidisciplinary team composed of parents, general and special educators, an educational administrator, an individual with expertise in interpreting assessments (usually a school psychologist), other service providers, others with special knowledge of the child, and the child (if appropriate). Parental consent is required at several points in the process, including at the initiation of the formal evaluation process. After the child is determined to have an educational disability, an IEP is developed. Once the IEP is implemented, the student's progress must be reviewed annually. Every 3 years, the student's continued eligibility for special education services is reevaluated.

Identification processes vary according to disability category. Sensory, physical, and neurological disabilities typically involve medical diagnoses, whereas developmental disabilities rely on multiple psychoeducational assessments to identify traits of the specific disabilities. Psychoeducational assessments are a function of schools and typically involve an individually administered IQ test, an educational achievement test, behavior rating scales, personality assessment, psychological processing tests (e.g., memory), social and developmental histories, and classroom observation. Assessment is conducted both to identify potential eligibility criteria and to rule out potential exclusionary factors (e.g., environmental and physical causes). For example, a student who is frequently absent from school might not be eligible for special education despite poor academic achievement unless it is also determined that the child has an educational disability over and above the problems associated with infrequent attendance. Most often, a school psychologist administers and interprets the psychological assessments. Frequently, a school social worker, occupational therapist, speech/language pathologist, and special education teacher are also involved in the assessment process.

Establishing reliable and valid criteria for the identification of learning disabilities, the largest special education category, has been the single most controversial issue in contemporary special education. The central criterion of the current definition of LD, an aptitude–achievement discrepancy, is particularly controversial because it has not proved to be a reliable way by which to determine the presence or absence of learning disabilities, distinguish students with learning disabilities from other low-performing students, or predict who will benefit from remediation. Furthermore, although the importance of early intervention is well established, the practical effect of using discrepancy formulas is that students must wait to fail because an aptitude–achievement gap requires time to develop. For these reasons, researchers and members of the PCESE have called for the development and use of responsiveness-to-instruction or responsiveness-to-intervention (RTI) models for identifying students with learning disabilities. RTI models avoid the "wait to fail" problems associated with current special education identification by providing progressively more specialized instruction to all students performing below local achievement norms prior to consideration for special education eligibility. Many students will respond to this increasingly intense instruction and can leave the RTI program. Those who do not respond to these multiple levels of intervention are subsequently considered for special education.

3.3. Placement and the Continuum of Services

Until recently, students with educational disabilities have been served primarily by special education teachers using special materials and instruction delivered in classes and programs designed for specific disabilities (e.g., all students with mental retardation receive instruction in a special classroom designed for students labeled as mentally retarded). Two recent trends have altered conceptualizations about how placement decisions are made and where special education services are provided. First, researchers have not found consistent or compelling evidence that placement is the decisive factor in students' academic progress, particularly for students with mild or moderate disabilities. Second, many parents and educators believe that inclusion in the general education setting is the preferred context for special education services.

In the United States, contemporary special education is best conceptualized as services provided to students regardless of location rather than as placement in specific programs (Fig. 2). Special education services are selected to meet student needs identified on the IEP. Although there are many variations, the large majority of students receive some combination of the following: (a) a resource room provided to supplement general education instruction, usually for a limited time of time (e.g., one class period) each day; (b) a special class provided to replace general education instruction for anything ranging from a single subject (e.g., reading) to the full day; and (c) a general education class with special education support ranging from consultation to in-class instruction by special educators. In addition to instructional services, many students in special education receive related services such as counseling, physical therapy, occupational therapy, and speech/language therapy.

Inclusion, the practice of educating children with disabilities in general education contexts, has a history of controversy. Early disputes were often ideologically based, positioning those who saw dedicated special education contexts as segregated environments with negative connotations against those who feared that students' individual needs would remain unmet in the general education context and that some special education students would disrupt general education learning environments. In most places, the inclusion

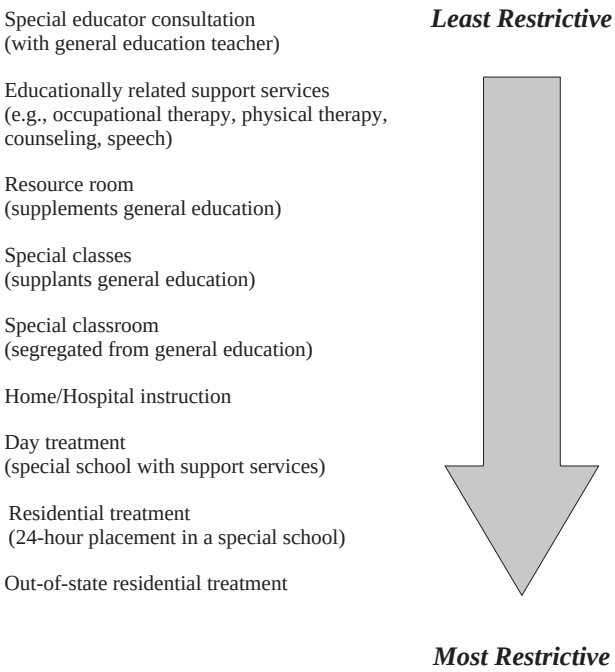


FIGURE 2 Typical continuum of special education services.

debate, benefiting from increasing experience and research, has moved to considerations of how to be successful in including all students in general education classrooms. Questions relating to providing special education services in general education contexts now address issues of teacher attitudes, training, and collaboration; supports and accommodations for students with disabilities, including materials and technology; effect of inclusion on other students (e.g., disruptive behavior, changes in curriculum or instructional materials); and costs (e.g., smaller class sizes, increased planning time for teachers, technology).

Most special educators now believe that placement decisions should be made on an individual basis rather than on a categorical basis because different educational contexts may address some learning and management needs better than do others. Where socialization or learning subject content information is the primary goal, a general education context might be most appropriate. Where a student's learning rate or level is highly unlike that of peers, a separate setting providing small group intensive instruction might be best. In any setting, effective teaching and consistently implemented instructional plans are critical for student achievement. Collaborative consultation between general educators and special educators, cooperative teaching arrangements, cooperative learning, peer-mediated instruction, peer tutoring, and direct instruction of skills all have been identified as effective educational strategies for children with disabilities in inclusion contexts. However, even when most instruction occurs in a special education setting, some degree of general education participation enables students with disabilities to be exposed to a wide range of academic content and to interact socially with nondisabled peers (e.g., having students with disabilities participate in all classroom activities but on a reduced basis with appropriate accommodations and modifications).

3.4. Instructional Strategies/ Effective Interventions

Many instructional strategies are available to meet the needs of students who require special education services. Some are effective for all students and can be used successfully in general education classrooms, whereas others have been developed to address unique learning needs of students with disabilities and are not likely to be used to address general education needs. In general, special education interventions fall into one of two broad categories: (a) interventions designed to remediate the disability or (b) interventions designed to compensate for the disability. For example, a student with a reading disability can (a) receive extra or different instruction in reading or (b) be provided with audio recordings of books. Both approaches have support, depending on the goal of the instruction. The following subsections briefly describe effective educational interventions for students with (a) sensory disabilities, (b) physical and neurological disabilities, and (c) developmental disabilities.

3.4.1. Sensory Disabilities

This category includes children with labels such as visually impaired, hearing impaired, and deaf/blind. Degrees of sensory impairment range from mild to profound. Thus, educational interventions include diverse approaches from adjusting classroom conditions (e.g., reducing environmental noise for students with impaired hearing) to providing specialized instruction and remediating secondary effects (e.g., teaching vocabulary and idioms to children who are deaf). Visual impairment refers to people who cannot see well even with correction. Low-vision devices and technology are used to support academic learning for some children with this disability. Educational interventions for students with more severe vision loss may include orientation and mobility skills. In general, students with sensory disabilities need to compensate for their diminished sensory input and might need to develop communication, vocational, self-help, advocacy, and recreational skills.

3.4.2. Physical and Neurological Disabilities

This category includes children with orthopedic impairments, other health impairments, traumatic brain injury, multiple disabilities, and autism. Students in this category have widely diverse educational needs. Some might lack the ability to walk or talk, whereas others might have a "hidden" disability such as a brain injury or another health condition (e.g., diabetes, asthma). These students might experience limitations due to the direct results of the disability or its effects (e.g., side effects of prescribed medications). In addition to modifications and adaptations in the learning context, teachers might need to monitor the physical or health conditions of these students. Students with physical and neurological disabilities, as well as those with sensory disabilities, often benefit from assistive technology to achieve academic success.

Historically, educational interventions for students with severe physical or cognitive disabilities have focused on social and functional skills required for survival in general society (e.g., communication, motor skills, decision making, functional reading, math) across multiple life domains, including vocational, leisure, home, and community settings. Instruction often focuses initially on discrete skills that are then linked to form complex behaviors. For example, a child might first learn to hold a toothbrush, then apply toothpaste, and then raise it to his or her mouth as part of learning the complex behavior of brushing one's teeth. Typical teaching strategies include verbal and physical prompting with eventual fading of prompts, praise for correct responses, instructive error correction, augmentative and alternative communication, and numerous opportunities to practice the new skill across multiple contexts.

3.4.3. Developmental Disabilities

Learning disabilities, speech and language disorders, and emotional and behavioral disorders are the three largest groups in this category. Learning disabilities are defined in the Individuals with Disabilities Education Act (IDEA) as "disorders in one or more of the basic psychological processes" that are evident in an "imperfect ability to listen, think, speak, read, write, spell, or do mathematical calculations." Some intervention strategies seek to improve academic outcomes by targeting underlying deficits in basic psychological processes. For example, a student may receive training in visual processing. However, to date, such indirect approaches have demonstrated little or no effect on academic achievement. Although it is possible that such approaches have merit, further development of accurate measures of basic psychological processes, as well as interventions that are designed to address these issues, will be required before such instruction is likely to be effective. In addition, although the modality-matched approach (e.g., programs for "visual learners") and the multisensory approach (e.g., using visual, auditory, and/or kinesthetic cues in instruction) are popular with many teachers, they have not been supported empirically.

In contrast, explicit, intensive, and supportive academic interventions have received substantive empirical support as successful methods for addressing academic failure. Effective instructional approaches include controlling task difficulty, using small interactive groups, teaching metacognitive strategies and problem-solving skills, using direct instruction, improving reading skills (e.g., phonemic awareness) and writing skills (e.g., organization and mechanics), and providing ongoing and systematic feedback. Such approaches improve academic achievement with students in both inclusion and special education classrooms, with students in a variety of disability categories, and with at-risk general education students.

Students with speech and language disorders, often referred to as communication disorders, may exhibit deficits in any of the following areas: receptive and expressive language, language development, articulation, fluency, and voice features. Speech and language pathologists provide a range of services in either private or classroom contexts to remediate these difficulties. Instruction may be targeted solely at learning the physical movements required for clear articulation (speech) or at developing language-related elements required for communication such as vocabulary and grammar.

Students with emotional or behavioral disorders may display inappropriately excessive or deficient behavior, possess unsatisfactory interpersonal relationships, and demonstrate academic learning problems. Intervention approaches used with these students include various psychotherapies, behavioral contingencies, family therapies, and segregation in residential programs. Behavioral and cognitive–behavioral approaches that include examining and altering antecedents and consequences associated with inappropriate behavior (e.g., functional behavioral assessments [FBAs]) and training in social skills have been shown to be effective in decreasing inappropriate behavior and increasing appropriate behavior. Because behavioral issues are central for most students labeled as emotionally disturbed, less research has been directed to identifying academic (vs behavioral) interventions specifically for students with emotional or behavioral disorders. However, instructional strategies that promote academic engagement and responding along with self-monitoring procedures have been shown to be effective. Because interventions for these students often require highly structured environments and more intensive instruction than is available in general education, this group of students is likely to receive direct services in a special classroom or school.

Students with emotional and behavioral disorders often become involved with the school disciplinary process. This can result in a conflict between the legitimate disciplinary needs of the school and the student's right to FAPE, particularly when the relevant behavior is severe or dangerous. In such cases, the school must conduct a "manifestation determination" to determine whether the behavior is a manifestation of the child's behavioral disability. If it is, then the child should not

be punished for manifestations of the disability. Instead, an FBA should be conducted to determine the situational and behavioral factors underlying the misconduct, and the student's IEP and/or placement should be changed to better meet the identified behavior management needs. If the misconduct is not a function of the student's disability, the student is subject to the normal disciplinary procedures.

3.5. Early Intervention and Adult Transition Services

Many educational disabilities can be addressed most effectively before children reach school age. Other disabilities are so severe that ongoing services are required beyond 21 years of age. U.S. federal legislation requires that states and local educational agencies provide special education and planning services to children who need intervention before 3 years of age and provide transition planning services to children who need adult services after 21 years of age. Such services provide early family-based interventions for young children and a planned transition to adult services for students who require support beyond the years covered by FAPE. Most often, early intervention and adult transition services are provided to children with moderate to severe sensory, physical/neurological, and developmental disabilities.

3.5.1. Early Intervention

Early identification of disabilities and subsequent intervention are crucial for enhancing development and learning, preventing the development of additional problems, and helping families adjust to the needs of children with disabilities. Early intervention is mandated by IDEA for very young children, from birth through 2 years of age, and refers to a comprehensive array of services that include educational interventions, health care, and social services (e.g., family counseling, parent classes) for families as well as for the affected children. A focus on the family and home environment is intended to (a) equip parents to participate in all aspects of decision making and administer interventions for their children and (b) locate instruction involving movement, language, self-care, and cognition skills in natural settings. To plan for and document interventions, an individualized family service plan (IFSP) is developed. The IFSP is similar to an IEP for school-age children, but the IFSP provides for more holistic goals, includes family needs in the intervention plan, and designs instruction around naturally occurring routines (e.g., bathing).

3.5.2. Transition to Adult Services

Transition to postschool settings and services is an important service provided by schools for the many students with disabilities who require supportive services after public schooling ends (i.e., 21 years of age). In the United States, transition planning begins a few years before a student's public education ends and continues until the transition is actually made. The transition planning process, which addresses academic, occupational, living, and social skill needs, is best conceptualized as an approach built on abilities, options, and self-determination rather than as a disability-focused process. Vocational education, work experience, tutoring, extracurricular group activities, parental support, and collaborations with community businesses, organizations, and agencies are important elements to address in successful transition planning.

4. MEDICAL AND TECHNOLOGICAL PROGRESS

Current and emerging information and medical technologies provide considerable hope for intervention for some disabilities and create considerable concern about ethical implications for others. Schools are required to provide assistive technology for students who need such devices to participate in and benefit from FAPE. Assistive technology refers to devices and services that enhance the performance of individuals with disabilities by enabling them to complete tasks more effectively, efficiently, and independently than would otherwise be possible. Assistive technology is used principally for aiding persons with physical and sensory impairments and includes items such as alternative keyboards, switches, Braille printers, mobility devices, adaptive feeding instruments, voice synthesizers to aid speech, and eye-controlled devices to help people without the use of their hands to accomplish everyday tasks. Assistive technology can also be used for students with high-incidence disabilities who require technologies such as speech recognition software to help compensate for writing difficulties, pagers or personal digital assistants to help with organizational skills, computers that translate text to speech, and calculators to assist with math learning.

Other technologies offer the promise of substantially augmenting current skills and providing alternative communication strategies to people with disabilities that have not been treatable previously. For example, electronic cochlear implants allow some deaf children to hear, and emerging electronic vision technologies may soon allow blind people to see. Some new medical technologies are controversial. For example, widespread use of cochlear implants could result in the disappearance of the vibrant and rich deaf culture that has its own language, customs, and values. Widespread implementation of genetic counseling and other medical techniques to screen for disabilities may allow for the possibility of correcting some disabilities before birth (in utero) but also have ethical implications ranging from whether parents will choose to have less than "perfect" children to judgments about the value of people with disabilities as contributing members of society.

5. MULTICULTURAL AND RACIAL ISSUES IN SPECIAL EDUCATION

Special education in the United States is grounded in civil rights legislation. The current system and legislation were constructed to ensure that all students have equal access to FAPE regardless of disability. However, issues of minority overrepresentation in special education arose almost immediately after the widespread implementation of special education services. Some advocates have seen special education as an inferior educational track that has been used to isolate and segregate students from minority cultures in substandard classrooms. Two recent trends make the already critical concerns about minority (racial and cultural) overrepresentation in special education even more imperative. First, the demographic profile of the United States is changing rapidly. By some predictions, the public school population will be more than 50% non-White by 2050, as it already is in six states. Second, academic achievement for students of color has been identified as a national concern. In general, public education in the United States has not served students from minority cultures well, particularly when such children are poor and live in urban and rural communities. However, the basic problem of poor school achievement for children from minority cultures is only tangentially related to special education. Poverty and deficient educational programs (general education and special education) in many urban and rural schools are issues for students from minority cultures primarily because these students are disproportionately represented in low-income groups. Placing large numbers of non-White students in special education cannot possibly address the fundamental issues of poor achievement. Interestingly, the issues of minority overrepresentation in special education are limited primarily to the judgmental educational disability categories. The incidence of severe disabilities is relatively stable across socioeconomic groups and cultures.

6. INTERNATIONAL PERSPECTIVES

Although special education is established in most developed nations, acceptance of people with disabilities and awareness of special education needs vary widely in other parts of the world. In 2001, the United Nations Education and Scientific Cultural Organization (UNESCO) estimated that 98% of children with disabilities in developing countries do not attend schools. Obviously, special education is not a priority in countries that do not provide general public education to their citizens. However, intergovernmental organizations such as UNESCO have advocated for education for all children. In 1994, the Salamanca Statement and Framework for Action on Special Needs Education was signed by delegates from 92 governments and 25 international organizations. The statement acknowledged that children with disabilities should be included in regular schools and that schools should make changes to "accommodate [the children] within a child-centered pedagogy capable of meeting [their] needs." As a result, some efforts to include children with disabilities in regular schools occurred worldwide.

Still, the provision of special education varies widely between and within nations. Countries such as Britain, Denmark, Ecuador, India, South Africa, and China have adopted changes in their education systems to include children with disabilities into regular classrooms. Although educational systems in some developing countries are moving toward inclusion, inclusive education is still viewed a luxury of First World countries. Currently, special schools in developing countries (e.g., China) are often affiliated with religious or private philanthropic organizations. In such cases, education for students with special needs depends on the availability of private services, the ability to access such services, and knowledge of their existence. However, in many rural areas of developing countries, livelihood is still dependent on manual labor. In these cases, literacy is not considered

to be essential and children who might be classified as having mild or moderate disabilities in other settings can lead normal lives and might not be diagnosed.

See Also the Following Articles

Behavioral Assessment in Schools ■ Learning ■ Learning Disabilities ■ Learning Styles and Approaches to Studying

Further Reading

Cook, B. G., & Schirmer, B. R. (Eds.). (2003). What is special about special education? [special issue]. *Journal of Special Education*, *37*(3).

Friend, M., & Bursuck, W. D. (2002). *Including students with special needs: A practical guide for classroom teachers* (3rd ed.). Boston: Allyn & Bacon.

Hallahan, D. P., & Kauffman, J. M. (2003). *Exceptional learners: Introduction to special education* (9th ed.). Boston: Allyn & Bacon.

Kohler, P. D., & Field, S. (2003). Transition-focused education: Foundation for the future. *Journal of Special Education*, *37*, 174–183.

MacMillan, D. L., & Reschly, D. J. (1998). Overrepresentation of minority students: The case for greater specificity or reconsideration of the variables examined. *Journal of Special Education*, *32*, 15–24.

Odom, S. L., & Woolery, M. (2003). A unified theory of practice in early intervention/early childhood special education. *Journal of Special Education*, *37*, 164–173.

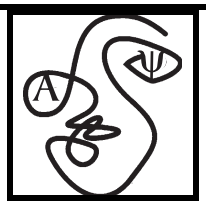
President's Commission on Excellence in Special Education. (2002). *New era: Revitalizing special education for children and their families*. Washington DC: U.S. Department of Education.

Vaughn, S., & Fuchs, L. S. (Eds.). (2003). Redefining learning disabilities as inadequate response to instruction: The promise and potential problems [special issue]. *Learning Disabilities Research & Practice*, *18*(3).

United Nations Education and Scientific Cultural Organization. (2001). *The flagship on education for all and the right to education for persons with disabilities: Towards inclusion*. www.unesco.org.

U.S. Congress. (1997). *Individuals with Disabilities Education Act amendments of 1997*. Washington, DC: U.S. Government Printing Office.

Witmer, L. (1907). Clinical psychology. *The Psychological Clinic*, *1*, 1–9.



Sport Psychology, Overview

Yuri L. Hanin
KIHU – Research Institute for Olympic Sports, Jyväskylä, Finland

Natalia B. Stambulova
Halmstad University, Halmstad, Sweden

1. Introduction
2. Competitive Sport as a High Achievement Setting
3. Performance Enhancement
4. Athletic Excellence from a Developmental Perspective
5. Conclusion: Future Directions in Sport Psychology
Further Reading

Sport psychology is a subdiscipline of psychology applied to a competitive sport as a specific context of organized physical (motor) activity. Competitive sport is focused on high achievement and consistent excellence, in contrast to other settings in which exercise is used for physical education, leisure, or rehabilitation. The major emphasis in sport psychology is on the study and application of psychological factors enhancing athletic performance and on the impact of sport participation on a person's (or team's) development.

GLOSSARY

athletic excellence An athlete's (or team's) exceptionally good performance process and outcomes compared with the previously achieved standards (self-referenced or normative). Athletic excellence reflects an athlete's (or team's) ability to perform consistently up to his or her potential by recruiting and utilizing effectively available resources to match the demands of a task at hand.

experiences The totality of past and present characteristics that makes up the particular quality of a person's performance. Performance-related experiences are reflected in situational states, relatively stable patterns (traits, qualities), and meta-experience (self-knowledge, attitudes, preferences/rejections of one's experiences).

resources Psychobiosocial (cognitive, affective, motivational, bodily, behavioral, operational, communicative) assets determining an athlete's ability to performance consistently up to his or her potential.

sport psychology A subdiscipline of psychology that deals with "psychological foundations, processes, and consequences of the psychological regulation of sport-related activities of one or several persons acting as the subject(s) of the activity" (FEPSAC Position Stand, 1995).

1. INTRODUCTION

1.1. Defining Sport Psychology

What is sport psychology? How is it different from other subdisciplines of psychology? How is it related to sport sciences? Although many definitions of sport psychology have been suggested, there has been no comprehensive and internationally accepted definition of sport psychology. In its Position Stand # 1 (1995), the European Federation of Sport Psychology (FEPSAC) proposed that "sport psychology is concerned with the psychological foundations, processes, and consequences of the psychological regulation of sport-related activities of one or several persons acting as the subject(s) of the activity" (p. 4). This definition indicates that sport psychology attempts to improve athletic performance and help athletes to concentrate better, deal effectively with competitive stress, and to

©2004 Elsevier Inc.
All rights reserved.

practice more efficiently. Moreover, sport psychology also attempts to understand the impact of long-term sport participation on development of personal resources of athletes in the setting of organized competitive sport. The term "sport" is used as an umbrella term that includes different kinds of sport, exercise, and other physically active pursuits. These types of physical activity are also used in other settings such as organized physical education, leisure, and rehabilitation (healing). Another important feature of sport psychology is its double nature. On the one hand, it is a part of psychology; on the other hand, its knowledge base is related to sport sciences focused on understanding human activity in this particular context. Thus, in applications, these two sources of knowledge help to better understand a person, the environment, and the key aspects of the sporting activity.

The major focus of this article is the context of competitive (high-achievement) sport. From this perspective, sport psychology examines mainly the short- and long-term impact of psychological factors on athletic performance and the potential effects of systematic participation (involvement) in sport. Applied sport psychology attempts to solve specific practical problems by improving athletic performance and thus helping athletes to develop their potential in the sport setting.

This article briefly reviews selected aspects of applied sport psychology within the framework of three basic constructs: athletic excellence, performance-related subjective experiences, and individual resources (psychological strengths). The key aspects of athletic performance are examined from the short-term (readiness for competition and performance excellence) and long-term (consistent excellence, career development) perspectives.

1.2. Major Focus and Trends in Sport Psychology

What are the major focuses in sport psychology research? What are the main trends in applied psychological work with athletes, teams, and coaches? Noteworthy are two major focuses in sport psychology research, with two corresponding trends in applied work. The first is understanding the psychological factors that affect athletic performance and how athletes realize their potential in sport. Applied aspects here include high-quality practices, optimal performance, and adequate recovery at the level of an individual athlete and team. The second important objective of sport psychology is to understand how athletes develop in sport and what are the "benefits" and "costs" of their multiyear sport participation. Applied aspects here include the need to help athletes cope successfully with career transitions and find a balance between sport and other spheres of life. In team sports, this also involves dealing with team-building issues and helping individual athletes find a balance between individual and team interests and values.

In competitive sport, applied psychologists deal with healthy, motivated, and high achievement-oriented people striving for consistent excellence, performance up to their potential, and continuous self-development. Thus, the focus on enhancement of athletic performance and empowering approach reflect a positive, proactive, and constructive nature of applied sport psychology. Interestingly, Seligman and Csikszentmihalyi (2000) called upon applied psychologists to move beyond studying psychological disorders and problems and spend greater efforts studying positive psychology that can be used to facilitate and enhance human functioning. This emphasis on positive psychology, or psychology of human resources and strengths, is not new, but has been occurring for the last 25 years. However, there is still an urgent need to attend to current concerns of athletes and coaches and examine more closely their successful experiences by bridging the gap between group-oriented and individualized approaches. Therefore, it is argued that sport psychology is the psychology of personal and athletic excellence and as such, from the very beginning was oriented to identifying a person's resources (strengths) to facilitate consistently successful performance up to the person's potential.

1.3. How Sport Psychologists Work

Who are applied sport psychologists? What do they do, and how and why do they work with athletes, teams, and coaches? These questions are important for an understanding of what sport psychologists can and can not do in the field of competitive sport.

First, sport psychologists as a professional group are experts with different backgrounds. They may be clinically oriented consultants, educationally oriented consultants, mental trainers, applied researchers specializing in performance enhancement, or social or personality psychologists. However, whatever their specialization, applied sport psychologists are usually required to be well versed not only in psychology

but also in sport and sport sciences. This helps them to establish and develop working relationships with individual athletes, teams, coaches, parents, managers, etc.

Second, it is well known that the science of coaching focuses on the use of general principles. Per Weinberg and Gould (1999), "the art of coaching is recognizing when and how to individualize these general principles" (p. 15). As with coaching, the practice of applied sport psychology is both a science and an art. As a science, it is based on various theoretical models and results of empirical studies describing what is typical for athletes in particular sport situations. As an art, sport psychology is grounded in the personality as well as personal and professional experiences of the consultant, and it is expressed in his or her ability to understand the particular athlete within a psychological context and to choose the most effective applied approach or intervention. That is why different consultants may work differently with the same athlete yet be equally successful.

Art and science aspects are sport psychologists' tools to help athletes and coaches, who often focus mainly on the symptoms or consequences of psychological problems, deal with real causes of the problems (challenges, task demands).

Third, there are certain organizational working models, assessment technologies, and interventions based on specific ethical norms that characterize how sport psychologists work. For instance, sport psychology research and effective delivery of psychological services to elite athletes and coaches usually focuses on two closely related aspects: (1) performance enhancement in practices and competitions, and (2) optimization of interpersonal and intragroup communication, creating optimal team climate and effective management. Sport psychologists use several guidelines or principles to enhance their work, including action and growth- orientation; an emphasis on developing individualized strengths rather than on repairs of deficiencies; empowering athletes, coaches, and teams rather than developing over-dependency on outside experts; and enhancing active participation, partnership, and cooperation between sport psychologists, athletes, and coaches.

In brief, working with an elite athlete or coach usually includes several action-oriented steps: (1) listening to the coach and athlete's account of the current situation and past performance history to identify their concerns that need to be addressed; (2) providing a general summary of how similar situations are usually handled in sport and suggesting a tentative plan of joint work on the problem at hand; (3) collecting the data and providing a detailed feedback with the interpretation of results using context-related language clear to the athlete and coach; (4) preparing an action plan for further analysis, change, and monitoring of the key parameters involved; (5) evaluating the effectiveness of the initial steps and developing an individualized intervention program with clear criteria to assess the athlete's progress on a daily, weekly, monthly, or seasonal basis; and (6) contacting (by phone, e-mail, or fax) the athlete and coach systematically, which is an important part of their work during the entire season. A wrap-up, "lessons learned" session is also a good way to summarize experiences of all participants by the end of the season. It is important to realize that this approach is different from the traditional role of an outside expert telling the client what to do or not to do. Sport psychologists' main task is to empower athletes and coaches via an individualized approach focusing on their strengths and successful experiences rather than on deficiencies and limitations.

2. COMPETITIVE SPORT AS A HIGH ACHIEVEMENT SETTING

2.1. Sport and Competition

What is an athletic competition? Why is it so important in sport? These questions relate to the psychological characteristics of sport and to social psychological contexts of competitions. Sport as a part of the world and national cultures is also a human activity in which people find, realize, and further develop their individual potentials. Organized competitive sport is characterized by a clear focus on high achievement, exceptional level of skills, enhanced working capacity supplemented by health, well-being concerns, and prevention of injuries. However, the key aspect in understanding the psychology of high achievement sport is competition as a social comparison process.

The essence of sport competition is an evaluation and a social comparison of athletes according to the specially developed and approved rules of that sport. Observable competitive performance is a process of delivering sport results by athletes or teams. Usually, judges measure results in competitions, and on the basis of the comparison between the participants, rank each athlete. In addition, athletes often use self-referenced (process- and/or outcome-oriented) criteria to interpret their results in terms of personal success or

failure. To demonstrate athletic excellence in competitions, athletes have to practice regularly and continuously develop their resources. Both practices and competitions contribute to the development of athletes' physical and mental competencies and skills required in their chosen sport event. However, only participation in competitions allows the athlete to demonstrate her excellence in public and thus to win social recognition and prestige.

Rules of competition in different sports create three distinct psychological contexts for competing athletes: (1) "one-by-one" performances (with no physical or psychological contact between opponents during performance); (2) "one-near-one" performances (with only psychological contact between opponents); and (3) "face-to-face" performances (with both physical and psychological contacts between the opponents during performance). Each of these contexts creates specific challenges for athletes and requires specific resources to cope with task demands. Moreover, a competitor can be either an individual athlete or a team. A sport team has specific structural and dynamic features, such as values, cohesion, communication, and leadership. Group processes can either expand or drain the individual resources of team members, thus affecting the quality of practices and achievement level in competition.

Since the beginning of the modern Olympic movement in 1896, sport has been developing immensely. Contemporary sport has become an international phenomenon and also a part of the world business. Increased mass media involvement has turned international competitions into prestigious social events where athletes often feel extremely high pressure from the social environment. The intensity of competition in high achievement and professional sports has increased dramatically, and in many sports the current level of results is close to the natural limits of human abilities. All this explains the increased role of psychological factors in contemporary sport, creating a challenge for applied sport psychologists to develop effective approaches in helping sport participants.

2.2. Individual and Team Excellence

What is athletic excellence? How is it related to the individual athlete and to the team? Is team excellence simply the sum of individual excellences? How do athletes in team sports find a balance between individual and team goals?

Athletic excellence is defined as an athlete's exceptionally good performance compared with the previously achieved standards. The standards of performance can be self-referenced, i.e., based on a particular athlete's record of achievements and performance history. In contrast, normative standards reflect performance levels of other top performers in a particular sport event. In both cases, the indicators of athletic excellence are results (outcomes) achieved and the quality of performance process (task execution). Athletic excellence is an indicator of athletes' ability to perform consistently up to their potential by recruiting and using effectively the available resources matching the task demands. On the other hand, the notion of personal excellence reflects a high level of ability to function effectively as a human being in different settings, including sport.

Depending on the type of sport activity, athletic excellence can be divided into individual excellence (demonstrated by an athlete) or team (or collective) excellence (achieved by a team). Although team excellence depends on individual contributions, it is often not equal to the sum of the individual performances. Therefore, a team composed of "star" athletes does not always demonstrate team excellence, whereas average players working for the team and sharing team values and high work morale can achieve outstanding team excellence. Research shows that team excellence requires not only individually outstanding performances but also adequate interpersonal and intragroup communication. These communication processes reflected in the team's values, norms, and leadership processes can provide substantial support for the unique resources of team members and compensate for the lack of other resources.

To achieve a collective excellence, it is important to find an adequate balance between the athletes' individual goals and the team goals. These goals usually overlap, but they often do not perfectly match. However, a coach should realize that the degree of this match or mismatch between individual and team goals can result in a balance or imbalance between cooperation and competition processes in the team. Specifically, higher overlap (a match) between individual and group goals leads to better cooperation between teammates, whereas a lower overlap (a mismatch) can result in competitive behavior among the players (e.g., competing for starting positions, playing time, etc.). To find an adequate balance between stimulating athletes to develop their individual excellences and encouraging them to contribute maximally to the team is one of the key issues for coaches.

Another important factor in developing a collective excellence is to identify individual resources and strengths of the players in order to give the players clearly formulated and interrelated roles as the components of specific tasks. Each task may be perceived as a challenge, a routine, or a risk, depending on the perceived relationships between the demands of the task and the available resources (individual and team). When resources and task demands match each other, the team has a set of challenges. Successful coping with challenges results in the development of available resources. However, if available resources exceed the task demands, the task may be perceived as too easy or routine, not requiring recruitment and effective use of resources. This may lead to boredom and low task involvement. Finally, if the task demands exceed available resources, then the task is perceived as a threat and a risk (of failure).

The distinction between challenges, routines, and risks is important for understanding the players' (and team's) development. Per Hendry and Kloep (2002), the lifespan model of developmental challenge states that development occurs when the "pool" of potential resources is added to and resources are strengthened. By contrast, stagnation describes a condition in which no new resources are added to the pool, or they are not strengthened. Finally, a developmental decay in an individual or a team performance is expected if the task demands exceed the potential (available) resources and thus drain the pool ceaselessly. The task of the coach is then to create specific challenges for an athlete (or a team) that will stimulate the effective recruitment, use, and development of existing resources (strengths).

2.3. High-Quality Practice

How much time do athletes have to spend in practice in order to achieve athletic excellence? What is the difference between high- and low-quality practice? The major focus in sport psychology since the late 1960s was on successful and poor performances in competitions. Although competitive stress is still a popular topic of research, it is clear that excellence in competitions depends on how much and how well athletes practice. Research shows that top performers typically engaged in 10,000 hours, or 10 years, of deliberate (effortful) and sometimes non-enjoyable preparation to become experts in their domain. Although becoming an expert does require a lot of work, practices must also be high-quality. Moreover, it is important to realize that there is a time limit of what is possible to achieve in quantity-oriented practice, whereas quality-oriented practice is limitless.

High-quality practices have several important features. First, they require an athlete to be very aware of his or her individual strengths and limitations, optimal emotional states, and bodily signals. An athlete should know how to recognize and monitor this working state during the entire practice and how to recover effectively. Moreover, each training session should have a special meaning for the athlete in relation to a long-term perspective of the season goals and specific tasks. One of the top Finnish alpine Skiers, Tanja Poutiainen, explained in a television interview the "secret" of her successful performances in the World Cup by emphasizing the role of high-quality practices. Specifically, she said, "Now I train differently. I focus thoroughly for each down hill race in practices. I know exactly what I want to achieve and I know what I am working on. It makes much more sense in what I'm doing now. Before I just did it, too often mechanically practicing different movement patterns." This athlete created a mind-set for a task that matched her resources, and these challenges helped her develop as an athlete and a competitor. She was able to learn more about herself and how to use her resources (skills) under different conditions from every practice.

Another important feature of high-quality practices is a simulation of specific competition conditions (e.g., time, competition rhythm, organization, track profile). Learning to focus on one's own game is another important characteristic of effective pre-competition simulations. If practices during the competitive season are more directly related to competition tasks, they serve as a more focused preparation for competitions. On the other hand, lessons learned in competitions provide useful ideas for more effective practices. Especially important are high-quality practices during a competitive season (training between and during several competitions). Basically, the focus of high quality practice is on recuperating, improving, and further developing one's physical, technical, tactical, and psychological resources. Such an approach is especially relevant in professional sport. For example, NHL ice hockey players usually play over 80 games during the season. The players do not have time for much practice, and it is not uncommon that the skills of these talented performers begin to deteriorate. Thus, consistent athletic excellence requires the conservation of available resources (physical, technical, tactical, and psychological strengths) through their recruitment, use, recuperation, and ongoing development.

3. PERFORMANCE ENHANCEMENT

3.1. Performance Related Experiences and Athletic Excellence

What is the difference between peak, optimal, and sub-standard performance? What are the optimal and dysfunctional experiences accompanying athletic performance? How do athletes develop competitive experiences?

As mentioned previously, athletic excellence is an extended period of exceptionally good performance by an athlete or a team that exceeds previously established or situationally acceptable self-referenced standards. The usual level of performance provides the frame of reference for defining individually successful (optimal, peak), less than successful (sub-standard, below average, plateaus), and poor (choking, slumps) performances. Peak performance describes an ideal (outstanding, desired) performance. In contrast, optimal performance is the greatest degree attained (or attainable) under implied or specified conditions (e.g., skill level, health status, opponents, weather conditions, competition site). Optimal performance is evaluated using the individualized (self-referenced) criteria based on an athlete's past performance history and present performance status. From this perspective, any athlete can attain an optimal performance, whatever her skill level.

Athletes' behaviors and subjective experiences accompany successful and less than successful performances. Pre-event emotional experiences affect performance, whereas ongoing performance affects the dynamics of mid- and post-event emotional experiences. There are three interdependent levels of human experiences related to and induced by athletic performance: (1) situational transitory emotional experiences (psychobiosocial states) such as anxiety, anger, joy, or excitement, (2) relatively stable patterns of experience (traits, dispositions), and (3) meta-experiences (experiences about experiences). For instance, an athlete can experience a high level of anxiety prior to a competition. This situational state manifests itself in negative thoughts and expectations, such as feeling nervous, worried, and apprehensive. This experience is very individual (idiosyncratic), and for different athletes it can be harmful, can be helpful, or may not affect athletic performance in a particular competition. If anxiety is experienced repeatedly, a consistent pattern of experiences or a typical response disposition (trait anxiety) is formed. However, the athlete often reflects on significant emotional experiences in particular situations and their effects upon athletic performance. As a result, meta-experiences are formed, and this self-knowledge, beliefs, and attitudes can strongly affect the athlete's interpretations of different performance situations and the choice of adequate (or inadequate) coping strategy.

For instance, Michael Johnson is often quoted as saying that "he was really nervous when he was not nervous prior to an important race." From previous experiences, he knew that high situational anxiety was an optimal experience for his performance. Specific meta-experiences usually trigger corresponding self-empowering or self-defeating thoughts and self-statements and thus determine the beneficial or detrimental impact of emotional state upon performance. Therefore, there is a special need for psychological help for athletes who are unaware of their optimal experiences or whose meta-experiences are less than effective (self-defeating).

There is ample research examining situational emotional states accompanying individually optimal (successful) and less than successful (poor) performances in different athletes across different and similar sports. For instance, the individual zones of optimal functioning (IZOF) model as an individual- and action-oriented framework developed in high achievement setting focuses on optimal and dysfunctional situational experiences accompanying both successful and poor performances. This individualized approach to description, prediction, and explanation of emotion–performance relationships employs a multidimensional conceptualization of emotion as a component of psychobiosocial state. The model predicts interindividual variability of emotion content and intensity and their effects on individual athletic performance based on the "in/out of the zone" principle. It is argued that different forms of psychobiosocial state (cognitive, affective, motivational, bodily, motor-behavioral, operational, and communicative) reflect availability of (or a lack of) resources, their recruitment and utilization, and a need for recovery (recuperation).

Briefly described, these findings indicate that (1) negative situational emotional experiences (such as anxiety or anger) are not always harmful for individual performance; (2) positive emotional experiences are not always helpful or optimal for performance; (3) optimal and dysfunctional emotional experiences are highly individual (idiosyncratic). An optimal emotional performance state is the one most favorable for a particular individual (or a team) under specified conditions, and usually results in an individually successful performance that is equal to or better than realistically expected.

Research also shows that, in contrast to an ideal performance state (flow state) triggered by outstanding performance, optimal emotional states can be positive and negative prior to, during, and after performance. Positive optimal states are experienced when an athlete's resources match well with current task demands; positive dysfunctional states reflect a routine performance situation in which resources are available but are neither recruited nor used properly. The task is perceived as too easy, which results in excessive (demotivational) satisfaction (leading to complacency and less involvement in the task) and even boredom. Negative optimal states (anger, anxiety) reflect a threat (or a risk) situation (task demands exceeding available resources) in which an athlete attempts to actively cope with this imbalance. Finally, negative dysfunctional states reflect a situation when an athlete is unable to cope, with task demands exceeding currently available resources. Repeated experiences related to unsuccessful performance (slumps) and a failure to recuperate existing resources could result in chronic staleness, overtraining, and burnout.

Sport psychology describes different aspects of performance-related situational experiences that actually characterize a state of readiness for competition. These include self-confidence (state and trait), attention and concentration, experiential and behavioral manifestations of intrinsic and extrinsic motivation at the situational and dispositional levels, and individually optimal levels of anxiety. This research indicates that the intensity of anxiety (as well as other emotions) associated with optimal sport performance varies considerably across athletes, even for those competing in the same event. It also indicates that a substantial percentage of athletes actually benefits from elevated anxiety; in these cases, interventions aimed at reducing anxiety may be counterproductive. These findings illustrate the notion that optimal anxiety reflects attempts to situationally compensate for an apparent lack of resources as related to task demands as a person-specific coping strategy. The effect is further enhanced by an optimal self-empowering meta-experience, i.e., the athlete knowing that such a level of anxiety is useful and helpful for him. The previously cited example of Michael Johnson illustrates this point.

On the other hand, athletic excellence requires an optimal and sustained effort, and the athlete's body must be appropriately energized, with physiological and psychological resources prepared for the stresses and physical demands of competition. Arousal as a component of the psychobiosocial state manifests itself in physiological reactivity and physical energy. It is also paired with varying levels of concomitant cognitive, affective-volitional, and motivational activity, and behavioral display.

Typically, the physiological component of arousal is measured through muscle tension, cortical activity, electro-dermal activity, respiration, and biochemical markers such as epinephrine and cortisol. However, in recent years, there have been numerous attempts to identify idiosyncratic markers of perceived subjective bodily response to competitive stress both in best and worst competitions. This line of research has a great potential for the practice of sport psychology by providing a tool to enhance an athlete's bodily awareness.

3.2. Resources as Performance Enhancement Strategies

What are the internal and external resources that can enhance athletic performance? How can athletes and teams use their resources more effectively?

The construct of internal and external resources proposed here is not entirely new. For example, it is used in the conservation of resources (COR) model proposed by Hobfoll to define and explain psychological stress. Examples of broadly defined resources include not only personal characteristics (self-esteem, mastery, well-being) but also interpersonal, material, and work-related resources. The basic tenet of the COR model is that people strive to retain, protect, and build resources because the potential or actual loss of these resources is a threat and a source of psychological stress. Then psychological stress is defined as a reaction to the environment in which there is (1) the threat of a net loss of resources, 2) the net loss of resources, or (3) a lack of resource gain following the investment of resources.

Hendry and Kloep proposed the lifespan model of developmental challenge, which employs the constructs of resources and challenges to explain the processes of human growth. Examples of potential resources include biological dispositions (health, personality, talents, intelligence, body shape, attractiveness); social resources (trust, attachment, size and quality of network); skills (basic, learning, social, psycho-motor, etc.); self-efficiency (self-efficacy appraisals, experience with success, assurance from others, locus of control); and structural resources (country, race, class, family, income, gender).

In competitive sport, resources are defined as psychobiosocial assets that determine athletes' ability to

consistently perform up to their potential. Here the emphasis is on how available resources are identified and then systematically and effectively recruited, used, recuperated, and further developed. There are four closely related approaches to the enhancement of athletic performance: situational and individually optimal states, relatively stable experience patterns (dispositions, personality traits, and sport-specific qualities), psychological skills, and group dynamics factors. The strategies used in each of these four approaches to performance enhancement are actually different groups of internal and external resources. There is considerable overlap between these four groups of resources, and they also concur well with the COR model and the lifespan model of developmental challenge.

Earlier discussion of individually optimal situational states and relatively stable experience patterns indicates that an athlete's awareness of her optimal states and adequate meta-experiences can be a very effective internal resource. Furthermore, personality characteristics and sports-specific qualities (e.g., winner's profiles, a wheel of excellence) are also important potential resources for achieving athletic excellence. Although personality characteristics do not directly predict situational performance, they could be instrumental in predicting long-term effects of sport participation, for example the mental health model proposed by Morgan (1985). However, it is important to realize that athletes can achieve success even if they lack certain personality traits and qualities. The implication for a consultant working with an athlete (or a team) is clear: he or she should focus on developing available individual strengths rather than repair apparent deficiencies.

Psychological skills as a set of techniques and coping strategies aiming to produce an optimal state of readiness are valuable resources that successful athletes learn and use systematically to achieve consistent excellence. These resources are usually targeted at some specific modality. Therefore, the classifications of psychological skills usually include implicit or explicit reference to some form of human functioning (e.g., cognitive, affective, motivational, bodily, motor-behavioral, operational, communicative). At the same time, these different forms of functioning as components of psychobiosocial state can also be employed to describe different task demands.

Group dynamics and environmental factors are also important personal and team resources. These potential resources include cohesion, psychological climate in the team, patterns of interpersonal communication among the teammates and between the coach and the players, leadership style, and group norms and values reflecting sport subculture. For instance, high work morale and honesty as accepted values and group norms in a sport team could promote fair play behaviors and considerably minimize cheating in sport.

From social psychological and environmental perspectives, it is important to realize that competitive sport is a part of the society at large. Therefore, equitable or non-equitable conduct found in society is generally reproduced in sport settings. When sport traverses racial, ethnic, social, national, and gender boundaries, it has the power to bring diverse people together while de-emphasizing social or cultural differences. In other words, fair treatment in sport, in contrast to other settings, can provide conditions that significantly extend existing personal (or team's) resources. As an external resource, while sport offers the opportunity to traverse cultural identities and unite different peoples, it also can have the opposite effect. Thus, unfair treatment or even discrimination in sport can overtax the athletes' and teams' resources and considerably slow down their situational success and long-term development. In some cases, however, this negative treatment can be a strong motivator for athletic and personal excellence.

As Krane argues, fair treatment in sport occurs only when there are equitable resources and opportunities for all participants, regardless of social group membership (e.g., gender, race, ethnicity, nationality, sexual orientation, and social class). In reality, that is not always the case. In some cases, differential treatment occurs, in which minority social group members are treated unfairly. In other situations, sport is an avenue for educating people and increasing awareness about a wide range of social issues, fighting social injustice, and providing humanitarian assistance. Therefore, applied sport psychologists may employ a variety of strategies to promote fair and equitable sport.

3.3. Barriers to Athletic Excellence

What are the barriers to optimal athletic performance? How can athletes (and teams) minimize or cope with them? The threat of a net loss of resources, the net loss of resources, or a lack of resource gain following the investment of resources can be strong barriers to successful performance.

Four groups of internal and external resources (situational states, personality traits, psychological skills, and group dynamics factors) proposed earlier

can provide a framework for describing potential barriers to athletic excellence. Specifically, the notion of resources and their role in enhancing athletic performance is dialectic. A lack of resources or a failure to identify, recruit, and use them effectively could become a potential serious barrier to consistent athletic excellence. Examples of such barriers are dysfunctional emotional states, an overemphasis on apparent deficiencies, and a lack of performance-related skills. Finally, environmental barriers include inadequate motivational climate in the team, selfish behaviors of teammates, media pressures, and conflicts between a coach and an athlete. The typical consequences of the impact of barriers include performance slumps, overtraining, burnout, and injuries.

To minimize detrimental effects of internal and external barriers, it is recommended that an athlete's awareness of available resources and the strategies of their ongoing development is enhanced. Such awareness should be extended to a better understanding of causes of sub-standard performance and learning better risk management by maintaining self-efficacy, emotional control, and individually effective coping skills. Although the emphasis here is on situationally effective coping strategies, their role should also be understood from a wider (career development) perspective.

4. ATHLETIC EXCELLENCE FROM A DEVELOPMENTAL PERSPECTIVE

4.1. Athletic Career Demands, Coping Resources, and Barriers

What do athletes have to go through in order to achieve athletic excellence? Can any athlete reach it? What factors help athletes to reach excellence, and what might act as barriers along the way? These questions relate to the athletes' development during their athletic career and the demands they have to cope with by using specific resources.

A rapidly growing body of research in sport psychology focuses on "athletic careers" in an attempt to better understand how different athletes in different sports become expert performers and how they reach and maintain consistent excellence. Metaphorically, the athletic career (from initiation of sport participation to the retirement from sport) can be described as a miniature lifespan course involving a number of important transitions between the predicted stages. Understanding the mechanisms of these transitions and stages is important for coaches, athletes, parents, and sport psychologists.

It is very common for an athlete taking the first steps in his sport to dream of reaching athletic excellence, turning professional, and winning the world championship or Olympic games. However, it usually takes a long time to make this dream true. A so-called "athletic pyramid" shows metaphorically that only a few athletes achieve athletic excellence and have successful (elite, recognized, professional) athletic careers. For instance, a pyramid with one professional soccer player at the top contains 6000 soccer players at bottom; a pyramid with one professional basketball player at the top has 14,000 players at the bottom. According to Csikszentmihalyi and Robinson (1986), "in highly competitive domains, such as music, math, or sports, the way down is always much broader than the way up. Year by year, it becomes more difficult to catch up, and dropping out becomes increasingly easy" (p. 275).

The athletic career of each individual athlete is unique, and there are still debates in sport psychology about factors contributing to individual differences in sport achievements. It is becoming increasingly clear that interplay of several groups of factors can help or hinder an athlete's development and achievement of athletic excellence. These factors include an athlete's innate talent/potential, environmental factors (competent coaches, family support, adequate conditions for practice, etc.), and an athlete's ability to develop, recruit, and use effectively all resources necessary to cope with the increasing demands of the athletic career.

An athletic career usually starts at the age of 7 to 10 years, sometimes even earlier depending on the sport event (e.g., in swimming, artistic gymnastics, figure skating, and ice-hockey). First, children perceive sport as merely "playing a game," however, later their attitudes change and sport becomes as "a sphere of education". Much later for those who reach the top it becomes a job or professional activity. It takes usually about 10 years of deliberate practice to reach an expert performance level in sports, and once there, the period before retirement usually lasts between 5 and 15 years. At the most general level, an athletic career typically consists of several stages: initiation, development, mastery/perfection/culmination, maintenance, and discontinuation.

Athletes striving for athletic excellence and staying at the top have to cope with increasingly complicated demands related to their practice, competitions, communication, and life outside of the sport. There are specific demands at the beginning (transitional) part of each athletic career stage. Research findings

summarized in the 1997 FEPSAC Position Stand # 3 indicate that the beginning of sport specialization requires adjustments to the demands of the sport event, coach, sport group, and a new schedule of everyday life. Young athletes must ensure the right choice of sport, show ability in learning sport skills, and test themselves in their first competitions. When the athlete and coach decide to work for results, they enter the transition to the development stage or intensive training in the chosen sport, characterized by more intense and specialized practice and participation in higher level competitions. This transition requires that athletes adjust to higher physical and psychological loads, improve their technical and tactical skills, achieve relatively stable results in competition, and balance the time and energy taken by their sport with other activities (studies, leisure, etc.).

The first significant success brings the athlete to top-level sport with its tough competitions, and that indicates a transition to high-achievement and "adult" sport, or to the mastery/perfection/culmination stage of athletic career. Further progress requires that athletes revise their lifestyle so that it works for their sport goals. They should also find their individual paths in sport, the ways to cope with pressures of selection to important competitions, and to gain respect of the team, opponents, and judges. In short, this stage is where an athlete earns her reputation, which later will work for her. Transition from amateur to professional sport is marked by adaptation to specific requirements and pressures of professional sports, competitions with very strong opponents, more independent training, and striving not only for the victory but also for fans' sympathies. The transition from the culmination to the maintenance stage of an athletic career is characterized by the necessity to search for additional resources in order to maintain a high level of achievements and to plan athletic retirement. The termination of athletic career is marked by leaving sports and transitioning to some other professional career, with adjustments to a new status, lifestyle, and social network.

The career demands briefly described above characterize so-called normative transitions. However, each athlete also experiences a number of non-normative (idiosyncratic) transitions related to his or her particular situation or environment. Transition demands create developmental conflict between "what the athlete is" and "what he or she wants or ought to be," which stimulates the athlete to find additional coping resources. The effectiveness of coping depends on the dynamic balance between transition resources and barriers. Transition resources includes all internal and external factors that facilitate the coping process (e.g., the athlete's self-knowledge, skills, personality traits, motivation, availability of social and/or financial support). Transition barriers include all internal and external factors that interfere with effective coping (e.g., a lack of necessary knowledge or skills, interpersonal conflicts, difficulties in combining sport and studies or work). Interestingly, the same experience may be either a resource or a barrier depending on the situation. For example, athletic identity, which according to Brewer *et al*. (1993) is "the degree to which the individual identifies herself with the athletic role" (p. 237), is usually an important resource for an athlete, especially when she is at the peak of her career. However, it can turn into a serious barrier in the process of adaptation to post-athletic career life.

Typically, at the beginning of their athletic career, athletes experience a lack of internal resources (sport-specific knowledge and skills), which are compensated for by social support from a coach, the family, and peers. At the culmination of their athletic career, athletes are usually at their most resourceful and their career demands are the highest. Elite athletes often rely very much on their relatively stable experience patterns and meta-experiences. At the maintenance stage of their athletic career, athletes often lack social support; their health deteriorates; they are bothered by the consequences of injuries, a lack of energy, and pressures in other spheres of life. However, all these concerns can be compensated for by the individualization of all aspects of the athletes' preparation. For example, veteran athletes typically train less than their younger counterparts, but they use their individual strengths more effectively. This allows them to maintain results at a high level until the very end of their athletic career.

4.2. Athletes' Successful Transitions and Crisis Transitions

What happen in a career transition when the athlete is either able or fails to cope with transition demands? Do athletes need any psychological help while in such a transition? If yes, what kind of help would be useful? These questions relate to the coping process, the outcomes, and consequences of career transitions.

The coping process is central in a transition and includes all strategies the athlete uses in order to adjust

to particular transition demands. An adequate match between the perceived demands and available resources creates a state of readiness to the career transition and a higher probability of successful transition. Successful transition is associated with effective coping when the athletes are able to recruit, use, or rapidly develop necessary resources and avoid (or overcome) potential transition barriers. One of the principles in effective coping is relying on athletes' strengths, which can compensate for potential and existing weaknesses or barriers.

An alternative outcome is a crisis transition, when an athlete is unable to cope effectively on his own with the demands of the transitional situation. Research identified a set of symptoms or markers describing typical reactions of athletes in crisis transition, including a decrease in self-esteem (as a first reaction to ineffective coping) and chronic emotional discomfort. Athletes also report new fears, increased sensitivity to failures, poor decision-making, and inadequate behaviors. Attempts to change the situation are usually ineffective, and instead of improvement new mistakes (and failures) are committed. Therefore, athletes in a crisis transition often describe feeling like they are in a blind alley or a dead end. For instance, as reported by Stambulova and Lindwall in 2002, one elite athlete who dropped out of sport after becoming caught up in doping described her feelings in the crisis as follows: "I totally panicked and did a terrible error . . . I took . . . forbidden substances as a final effort to get away from the feeling of being useless. My head was in chaos and there were no open roads left to take" (*Svensk Idrottspsykol.*, 2, 2–5).

Athletes in crises need psychological assistance to shift them from a "dead end" situation to a "cross-road" situation and to see several new coping alternatives. Moreover, psychological intervention influences the consequences of the transition. Effective intervention leads to successful but delayed transition. Alternatively, ineffective or no intervention situations are followed by negative consequences or so-called costs for failure to cope with the transition. Possible costs include decline in sport results, injuries, overtraining, neuroses, psychosomatic illnesses, prematurely quitting sports, and also different forms of rules violation and degradation of personality (e.g., alcohol and drug use, criminal behaviors). All these costs can be seen as negative effects of sport participation and also as barriers to coping with forthcoming career demands.

A developmental perspective provides a framework for a better understanding of career transitions. For instance, Vygotsky's constructs of the zone of actual development (ZAD) and the zone of proximal development (ZPD) could be instrumental in prediction of transition consequences. The ZAD is a range of the tasks that a person can solve on her own; the ZPD is a range of the tasks that a person can solve only if assisted by others. If most of the athlete's coping resources are in her ZAD, a successful transition can be predicted; in contrast, a crisis transition is expected if most of the athlete's resources are in her ZPD. Therefore, a psychological intervention should focus on helping an athlete to develop new resources and overcome potential transition barriers, especially if transition demands exceed available resources.

The lifespan model of developmental challenge earlier applied to athletic performance (from a short-term perspective) can also be used for the interpretation of career transitions. Typically, transition demands require a long-term coping process and many resources. Successful coping means adding new resources and an outcome in the form of development. If no new demands are made and the athlete simply repeats everyday routines, development eventually turns into positive stagnation. Crisis transition can be seen as negative stagnation, which might turn into development (under condition of qualified psychological assistance to the athlete) or into decay (i.e., negative consequences of not coping with the transition).

Developmental psychology interpretations demonstrate the dialectic nature of career transitions and their role in achieving athletic excellence. Each career transition with its accompanying demands is a step to athletic excellence. There is a risk of not meeting the demands, resulting in negative stagnation or decay. But there is also a chance to develop further and to experience positive stagnation on a higher level.

Three types of psychological interventions can be useful for helping athletes in career transitions: (1) a crisis prevention, (2) psychological crisis coping, and (3) psychotherapeutic (clinical) interventions. Crisis prevention involves career planning and goal setting, mental skills training, and organization of a social support system. This intervention aims to prepare athletes for a transition in advance by developing their resources for effective coping. This approach actually enhances their readiness for the transition either on their own or by using an expert assistance. Psychological crisis coping is an intervention for athletes already in crisis transition; it includes mainly individual counseling and psychocorrection programs. The focus here is on helping the athlete to analyze her situation, to find the best option for coping, and to develop and realize the action plan. These interventions usually deal with negative

stagnation and help the athlete to turn it into a development situation. Psychotherapeutic or clinical interventions are applied when the athlete has already experienced one or several of the above-mentioned negative consequences of not coping with a crisis transition. In other words, these interventions deal with a decay situation, trying to stabilize and then to improve the athlete's situation.

4.3. From Athletic to Personal Excellence

What are the benefits and costs of many years of participation in sport? How can sport psychologists help athletes to maximize the benefits of an athletic career and to minimize its costs? How can a successful athletic career contribute to the athlete's life outside sport?

An athletic career can be evaluated not only as a stage-like developmental process, but also as a developmental event contributing to the lifespan development in and outside sport. From this perspective, several parameters characterize an athlete's development during his or her athletic career. These include duration of sport participation from start to peak and finish, the sport event(s) practiced, the degree of specialization, and achieved sport titles/records/results. Subjective indicators include perceived benefits of sport participation and its costs (in terms of time, energy, health, money, etc.) as well as career satisfaction (one's self-esteem in regard to the athletic career) and career successfulness (social recognition of one's athletic career).

Successful (or elite) careers are usually associated with athletic excellence, whereas satisfactory careers are associated with achieving individual peaks corresponding to the individual resources and environment. Satisfaction is based on a set of self-referenced criteria, but most often it consists of perceived potential in relation to level of achievements and athletic career costs. Interestingly, some athletes are often satisfied with non-elite careers, if they value the developmental effects (benefits) of sport participation (e.g., self-knowledge, physical fitness, good health, skills, qualities, social contacts that can be used in other spheres of life). In contrast, other athletes may be dissatisfied with their elite careers, especially if they perceive the costs as too high (e.g., deteriorated heath, deficits in education, a lack of close personal contacts or any interests outside sports).

To achieve athletic excellence, athletes have to start and specialize in a particular sport event quite early. This can facilitate young athletes' progress in a chosen sport, but it also can result in several negative consequences such as high pressures, fears, and one-sided development. To avoid this, coach effectiveness training encourages coaches working with children and youth to focus more on optimal development of young athletes than on "winning at all costs." Positive developmental effects related to athletes' self-esteem, skills level, and satisfaction with various aspects of sport participation should be provided for all young athletes, and then allowing the most talented of them to move further to the athletic excellence level.

In a broader sense, sport psychology aims to help all athletes, including top performers facing tough transitions and pressures of their careers, to achieve optimal development and their individual peaks in sport. Therefore, career/developmental perspective in applied work with athletes includes several aspects: (1) "whole career" approach, which spans the athletic career—from initiation to termination—as well as the post-athletic career; (2) "whole person" approach (taking into account not only athletic but also non-athletic developments of athletes); (3) developmental approach (links between past, present, and future); (4) activity-specific approach (taking into account general and sport event-specific factors); (5) individual approach (taking into account typical and individual patterns); and (6) transferable skills approach. For instance, the latter refers to a series of sport-based life skills programs that aim to teach physical and mental skills (e.g., emotion self-regulation, effective communication, goal setting, coping with success and failure), which can be generalized to various spheres of the participants' life outside sport. This approach can be useful at each stage of athletic career, especially for retired athletes, to help them adapt their skills and experiences acquired during sport participation to their post-athletic career life.

A challenge for sport psychologists helping athletes to reach athletic and personal excellence is to find the right balance between situational current problems and future career development issues. For example, what is more important for the athlete: to prepare well for a competition in the next week or to make sure she joins a national team in the next year? The other dilemma, for example, with a veteran athlete, is whether to focus on searching for additional resources to help him keep his sport results at a high level or to plan for retirement and post-athletic career life. The best answers to these and other similar questions can be provided by viewing applied sport psychology as both a science and

an art. The science viewpoint tells us that it is important to keep in mind both the situation and career perspectives; whereas the art viewpoint, based on past experiences, skills, and intuition, can help answer the question of how to do this.

5. CONCLUSION: FUTURE DIRECTIONS IN SPORT PSYCHOLOGY

What has been achieved in applied sport psychology? What are the main concerns of the field right now? What is on its future agenda?

In order to enhance the effectiveness of scientific support in elite sport, several new future directions from a research-oriented and a practical (organizational) perspective can be identified. These include a new emphases on the role of elite coaches in psychological preparation of athletes and team, more focus on team-building, environmental, and organizational factors, and the development of closer international cooperation between scientists, practitioners, and sports organizers. Each of these aspects is briefly described in the sections that follow.

More psychological support for elite coaches. Initial focus of most sport psychology research and interventions on athletes and teams is well documented in the literature. However, the role of coaches in the psychological preparation of athletes and teams should be further emphasized. In practice, this means that the coach should be the central figure in preparation of the team, and sport psychologists should work more through the coach and with the coach-athlete team rather than only with the athlete. Enhancing the psychological competence of coaches can be a decisive factor in enhancing the quality of coaching.

In the past, sport psychology interventions and mental training programs usually focused on competing athletes who were coping with competition stressors. Less attention was paid to high-quality practices and prevention of overtraining, staleness, burnout, and injuries. Therefore, a most urgent and promising area of research and applications in sport psychology now and in the future should be the optimal performance of coaches and their coping skills for handling short-term and long-term chronic (e.g., burn-out) stresses. Qualitative research on careers of outstanding coaches to identify the factors of their consistent excellence would be a challenge for future researchers and practitioners. On the applied side, it would be helpful to summarize experiences of how ongoing individualized consultancy (personal coaching) for coaches has helped them anticipate the critical transition periods in their careers.

Team-building and effective management. In the past, social psychological research in sport psychology comprised 8–10% of all efforts, and the role of environmental and organizational factors in elite sport is still underestimated. Therefore, sport psychology should focus more on a holistic approach to the interpersonal and group processes that are determining performance and the life of a team in a wider social and cross-cultural context. Optimization of communication in the team is a very promising and productive area of research and applications. Practically, very little is known about the psychology of effective management in elite teams, sports federations, and clubs. Considering the quick development of elite sports, such areas as organizational development, change, and change management are potentially very important as new directions for research and applications. Experiences and practices of organizational psychology and management already available in non-sports high achievement settings could be beneficial for sport. On the other hand, the findings obtained in elite sports might be of interest to top management, business, army, and police.

Cross-cultural adaptation of athletes and coaches. Recent developments in Europe and worldwide indicate that more and more elite athletes and coaches are working abroad. These professionals need new skills for successful adaptation to a new environment and its constant changes. Quick adaptation to a new team, teammates, and coach, effective contacts with the media, and negotiation skills, for example, are much-needed resources for elite athletes and coaches. Moreover, with more migration and higher mobility rates among elite coaches, a critical factor is the assessment of a candidate's potential for cross-cultural adaptation and individualized programs that could facilitate his or her entry to a host country. This is especially important in view of the fact that tradition and values vary by country, and, for instance, a well-meaning but authoritarian coach with a clear orientation on success can be less than effective when starting his work in an amateur-oriented environment of the host country. A follow up with the coach or athlete could be instrumental in helping them to quickly adapt and effectively function both professionally and personally in the new environment.

International cooperation of sport psychologists. There are indications that in the future, a better collaboration

between applied sports psychologists from different countries could be useful not only for research but also in consulting. With recent developments in world-wide communication, joint consulting and psychological support for coaches and athletes across different countries seems like a reality in the near future. Developing such a network of sport psychologists could be an interesting initiative, especially in places where there is a lack of experts who could provide high-quality services (in research and applications) for elite athletes and coaches. One possible solution would be to use the expertise of internationally recognized applied researchers and practitioners in sport psychology who could deliver the necessary services for elite athletes, teams, and coaches and provide hands-on experiences for the local young aspiring sport psychologists interested in working with elite performers.

To conclude, now as never before, the application of what is already available in sport psychology is extremely important. Practical experience and expertise available in sport psychology are important not only in competitive and elite sport settings but also in such high-achievement settings as the performing arts and business. There are promising indications that the gap between theoretical knowledge and experience-based knowledge in sport psychology is gradually being bridged. Moreover, there is a clear shift in applied sport psychology from a predominantly negative, problem-oriented, and deficit-repairing approach initially borrowed from clinical psychology to a more positive psychology focusing on optimal performance and on an athlete's and team's strengths rather than limitations. Another promising trend in sport psychology is more emphasis on idiographic (individual-oriented) and experience-based approaches rather than on traditional nomothetic (group-oriented) comparisons of successful and less than successful athletes. Finally, early attempts to use personality tests to predict situational performance proved to be unsuccessful. A new and more promising approach is to conceptualize the situationally oriented applied work focused on enhancement of athletic performance within the framework of developmental perspective. This may provide an opportunity for sport psychology to become the psychology of athletic and personal excellence.

See Also the Following Articles

Anxiety and Optimal Athletic Performance ■ Arousal in Sport ■ Assessment in Sport Psychology ■ Attention and Concentration Training in Sport ■ Competition in Sport ■ Decision Making in Sport ■ Emotion in Sport: An Individualized Approach ■ Fair Treatment and Discrimination in Sport ■ Goal Setting and Achievement Motivation in Sport ■ Group Dynamics in Sport ■ Injury in Sport ■ Intrinsic and Extrinsic Motivation in Sport ■ Overtraining and Burnout in Sports ■ Performance Slumps in Sport: Prevention and Coping ■ Psychological Skills Training in Sport ■ Self-Confidence in Athletes ■ Successful Athletic Careers

Further Reading

Aspinwall, L. G., & Staudinger, U. M. (Eds.). (2003). *A psychology of human strengths. fundamental questions and future directions for a positive psychology*. Washington, DC: American Psychological Association.

Brewer, B. W., Van Raalte, J., & Linder, D. E. (1993). Athletic identity: Hercules' muscles or Achilles heel? *International Journal of Sport Psychology*, *24*, 237–254.

Csikszentmihalyi, M., & Robinson, R. E. (1986). Culture, time and development of talent. In R. J. Sternberg, & J. E. Davidson (Eds.), *Conceptions of giftedness* (pp. 264–284). New York: Cambridge University Press.

Danish, S. J., Petitpas, A. J., & Hale, B. D. (1993). Live development interventions for athletes: Life skills through sports. *The Counseling Psychologist*, *21*, 352–385.

Definition of sport psychology. (1995). FEPSAC Position Stand # 1. *FEPSAC Bulletin*, *7*, 3–5.

Durand-Bush, N., & Salmela, J. (2001). The development of talent in sport. In R. Singer, H. Hausenblas, & C. Janelle (Eds.), *Handbook of sport psychology* (2nd ed., pp. 269–289). New York: John Wiley & Sons, Inc.

Ericsson, K. A. (1996). *The road to excellence: The acquisition of expert performance in the art and sciences, sports and games*. Mahwah, NJ: Erlbaum.

Hanin, Y. L. (2003). Performance related emotional states in sport: A qualitative analysis. *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research*, *4*(1). Available at http://www.qualitative-research.net/fqs-texte/1-03/1-03hanin-e.htm.

Hanin, Y. L., & Stambulova, N. B. (2002). Metaphoric description of performance states: An application of the IZOF model. *The Sport Psychologist*, *16*, 396–415.

Hanin, Y. L. (Ed.). (2000). *Emotions in sport*. Champaign, IL: Human Kinetics.

Hanin, Y. L. (1999). Sport psychology research and consultancy with elite athletes and coaches. In E. Muller, F. Ludescher, & G. Zallinger (Eds.), *Science in elite sport,* (pp. 129–145). London, UK: E & FN Spon/Routledge.

Hendry, L. B., & Kloep, M. (2002). *Lifespan development: Resources, challenges, and risks*. London, UK: Thomson Learning.

Hobfoll, S. E. (1989). Conservation of resources: A new attempt at conceptualizing stress. *American Psychologist*, *44*, 513–524.

Orlick, T. (1990). *In pursuit of excellence. How to win in sport and life through mental training* (2nd ed.). Champaign, IL: Leisure Press.

Smith, R., & Smoll, F. L. (1991). Behavioral research and interventions in youth sports. *Behavior Therapy*, *22*, 329–344.

Smith, R., Smoll, F. L., & Curtis, B. (1979). Coach effectiveness training: A cognitive-behavioral approach to enhancing relationship skills in youth sport coaches. *Journal of Sport Psychology*, *1*, 59–75.

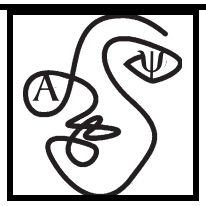
Sports career transitions. (1997). FEPSAC Position Stand # 3. *FEPSAC Bulletin*, *9*, 6–8.

Stambulova, N. B. (1994). Developmental sports career investigations in Russia: A post-perestroika analysis. *The Sport Psychologist*, *8*, 221–237.

Stambulova, N. B. (2000). Athletes' crises: A developmental perspective. *The International Journal of Sport Psychology*, *31*, *4,* 584–601.

Stambulova, N. (2003). Symptoms of a crisis-transition: A grounded theory study. In *SIPF yearbook* (pp. 97–109). Örebro, Sweden: Örebro University Press.

Vygotsky, L. S. (1984). *Sobranie sochinenii,* T.4 [The complete works, V.4]. Moscow: Pedagogika.



Stereotypes

David L. Hamilton and Sara A. Crump
University of California, Santa Barbara, California, USA

1. Introduction
2. Conceptual Approaches to Understanding Stereotypes
3. Stereotype Activation
4. Measuring Stereotypes
5. When Do We Use Stereotypes?
6. Consequences of Stereotyping
7. Stereotype Change
8. Conclusion
Further Reading

GLOSSARY

authoritarian personality Stereotypes and prejudice are thought to be the result of a personality syndrome that results from strict parenting style and is characterized by obedience to authority figures, strict adherence to rules, and intolerance of deviation from conventional norms.

categorization The mental process of organizing people, objects, and events into meaningful groups.

contact hypothesis The hypothesis that bringing people together in a supportive environment may help to reduce stereotyping and prejudice by fostering mutual understanding between members of different groups.

out-group homogeneity effect The perception that members of one's out-group are more similar to each other than are members of one's in-group.

priming An observed stimulus consciously or nonconsciously activates a related concept in the mind, which can then influence subsequent information processing.

self-fulfilling prophesy A person's prior expectancy (e.g., impression or stereotype) about another person influences his or her behavior and leads the other person to behave in ways that confirm the original expectation.

social desirability The desire to act in ways that are socially acceptable, even if the behavior does not reflect one's true beliefs.

subtyping The mental act of placing individuals who violate group stereotypes into their own separate category, thereby maintaining the original stereotype.

Stereotypes are cognitive structures containing our knowledge, beliefs, and expectancies about social groups. These belief systems arise from personal experience with groups and their members, through the transmission of social norms, and from people's tendency to organize and understand the world through categorization processes. Once formed, stereotypes have a number of consequences for people's perceptions and judgments of others as well as for their behavior.

1. INTRODUCTION

Throughout our daily lives, we encounter a vast array of people and situations that we perceive and interact with on a regular basis. Often, we perceive those people in terms of their memberships in some groups, and in many cases we have prior beliefs—stereotypes—about those groups. Stereotypes can then shape the way we perceive members of those groups, and they may influence our feelings and behavior toward them.

©2004 Elsevier Inc.
All rights reserved.

In 1922, journalist Walter Lippman coined the term stereotype to refer to people's penchant for consistently describing members of a group in the same ways. He conceived of stereotypes as "pictures in people's heads" that guide their perceptions of groups. More contemporary writers, building on Lippman's thinking, define stereotypes as people's beliefs about a particular group or social category. Specifically, stereotypes consist of a person's knowledge, beliefs, and expectancies about social groups. For example, we may believe that Italians are extraverted and passionate, whereas the Chinese are more polite and reserved. As these examples illustrate, stereotypes—even when based on some "kernel of truth"—are overgeneralizations in that they express the belief that all members of a social category share similar stereotypical characteristics. These beliefs may then influence our feelings about members of the group (prejudice) as well as our behavior toward them (discrimination). In short, whether accurate representations or not, we all use stereotypes to help guide our thoughts and expectancies toward others.

2. CONCEPTUAL APPROACHES TO UNDERSTANDING STEREOTYPES

Three major conceptual orientations have guided research on how and why people develop and use stereotypes: the psychodynamic/motivational, sociocultural, and cognitive approaches.

2.1. The Psychodynamic/ Motivational Approach

The psychodynamic/motivational approach views stereotyping as serving the needs of the individual perceiver, protecting and enhancing the self by derogating others. There have been several variations on this theme.

In early work guided by Freudian theorizing, stereotypes were seen as a response to unacceptable impulses or tensions within the individual. For example, when an individual feels that he or she possesses certain negative qualities, rather than directly addressing those issues, the individual may do so indirectly by projecting his or her own negative characteristics onto others. When those attributes are projected onto an entire group of people, the result is stereotyping. This process serves the function of avoiding recognition of those qualities in oneself. Reflecting this approach, in 1950 Adorno *et al.* proposed that stereotypes and prejudice arise from strict child-rearing practices that result in an "authoritarian personality." This personality syndrome is characterized by strict obedience to authority, rigid adherence to rules, and intolerance of deviation from conventional norms.

Recent research has focused on an individual's motivation to maintain a positive self-concept and its role in intergroup perceptions. For example, Tajfel and Turner's 1986 Social Identity Theory posits that people derive part of their self-esteem from their group memberships. Because people want to feel good about themselves, they therefore are motivated to feel good about the groups to which they belong. Consequently, people often derogate out-groups in order to make their own groups appear more favorable, thereby contributing to their self-esteem. In support of this view, some studies have shown that when a person's self-esteem is threatened, stereotyping of out-group members leads to an increase in self-esteem.

Finally, stereotypes may also help us to justify and perpetuate the status quo. For example, placing blame on those in our society who are disadvantaged (e.g., believing that African Americans are poor and of lower status because they are lazy) not only can help us to relieve guilt and maintain a positive self-image but also can perpetuate the disadvantaged status of less fortunate groups.

According to the psychodynamic/motivational perspective, then, stereotypes help us to deal with internal conflict and to maintain self-esteem.

2.2. The Sociocultural Approach

The second major approach to understanding stereotypes focuses on the influence of one's social and cultural environment on stereotype development and use. Specifically, according to the sociocultural approach, stereotypes are passed on to children through socialization processes and reflect the norms of society. At an early age, children become aware of the values held in their society through the teachings of their parents, in their interactions with their teachers and peers at school, and by what they see on television. For example, children may overhear their parents making derogatory comments about people of a different race or religion, their teachers may treat them differently depending on whether they are male or female, their peers may pick on certain children because they are developmentally delayed, or they may see members of some groups consistently presented in a stereotypical way on television programs.

In addition, certain groups of people are overrepresented in some roles in our society, potentially leading

children to believe that those positions reflect the inner dispositions of those group members. For example, children may observe that members of ethnic minorities are primarily employed in the lower paying jobs in the workforce, which may lead them to believe that those minorities are not hard workers or not intelligent enough to get better jobs. They may observe that men occupy most executive and leadership positions, fostering the belief that males are more aggressive and powerful than females. In each of these situations, children are learning the stereotypes held in their society by observing the distribution of group members in different social roles. Although that distribution may be arbitrary or may be constrained by external factors, observers may infer that it reflects internal, psychological characteristics, which then become part of the stereotype.

2.3. The Cognitive Approach

According to the cognitive approach, stereotypes develop in part as a result of people's natural tendency to categorize the people and events they encounter in the complex world in which they live. In fact, research guided by this approach has shown that the categorization of people into groups is a natural, fast, and effortless process that is used by everyone. For example, when meeting people for the first time we automatically categorize them in terms of their age, race, and gender—physical properties that are visually salient and are often believed to provide more information about a person than other physical features (e.g., brown hair). Social categorization can also be based on other features, including nationality, religion, occupation, sexual orientation, and many other characteristics. Through the categorization of individuals into groups, perceivers are able to more efficiently organize their knowledge about the world.

Categorization by itself does not mean that stereotypes are formed. However, once we differentiate people into different groups, we quickly begin to notice how those groups differ, and we form beliefs about the attributes that are characteristic of each group. Those beliefs, once formed, comprise the stereotype of the group. In this way, stereotyping derives from the foundation provided by the categorization process.

A related function of this categorization process is that the resulting stereotypes provide us with useful information. When encountering a member of a stereotyped group, the perceiver knows little about the individual. Yet by drawing on the stereotype, the perceiver can go beyond the information literally available and can infer a variety of attributes that are then assumed to be true of the person, given his or her group membership. In this way, the perceiver has expanded his or her "knowledge" of the person. However, this process also entails the loss of information because in perceiving this person as being typical of members of the stereotyped group, the perceiver ignores information about this particular individual. Thus, through mental shortcuts, categorization and stereotyping allow us to preserve cognitive resources for other activities but they do so at potential cost.

In summary, each of the three conceptual orientations discussed previously approaches the study of stereotypes from a different perspective. Despite their different emphases and foci, elements of each approach undoubtedly contribute to stereotype formation and use.

3. STEREOTYPE ACTIVATION

In our everyday lives we often use stereotypes in making inferences and judgments about others. What factors influence whether stereotypes will be activated in the first place? In many cases, stereotypes may be activated simply when one encounters a member of a stereotyped group. In other words, the encounter may make certain beliefs more accessible in our minds, and these beliefs may then influence our perceptions of the person. For example, meeting an old person at a crosswalk might activate the stereotype "slow," whereas meeting an old person teaching on a college campus might activate the stereotype "wise." Also, there are situations in which one's membership in a group becomes salient. For example, the only woman on a committee of 12 people becomes salient, and that salience can make stereotypes about gender especially prevalent in that situation. Also, for some people stereotypes about certain groups can become chronically accessible so that the stereotype is routinely used in judging group members. Finally, research has shown that stereotypes may be activated automatically, even without the perceiver's awareness, and then can influence judgments and behaviors. Thus, stereotypes can be activated in a variety of ways, and in all cases they are likely to guide subsequent perceptions and interactions.

4. MEASURING STEREOTYPES

Research on stereotype measurement began in a classic study by Katz and Braly in 1933. Participants were given a lengthy list of trait adjectives and for each of

several groups (racial, national, and religious) were asked to select all of the words that were characteristic of the group. The adjectives that were most frequently selected were interpreted as representing the stereotype of that group. Follow-up studies using the same method have been conducted to assess the degree of stability or change of various stereotypes. This research has found, for example, that the characteristics selected as descriptive of African Americans have become less negative over the years. Does this mean that the stereotype has changed such that Americans now view African Americans more positively?

Perhaps this is so, but, unfortunately, it is difficult to draw such a conclusion with certainty. Although many researchers have used Katz and Braly's method over the years, it has some flaws that limit its validity. One of the most obvious limitations is that the checklist measures people's self-reports, and these self-reports are subject to bias. For example, when people report their beliefs and opinions, it is unclear whether they are indicating what they privately believe about African Americans, what other people think about African Americans, or what they believe is socially appropriate to say about African Americans. One explanation for the finding that negative stereotypes toward African Americans have decreased is that people are being influenced by social desirability concerns (i.e., they do not want to appear racist). To the extent that such concerns influence these measures, they do not provide accurate measures of people's stereotypes.

Recent studies measuring stereotype content have attempted to overcome some of these limitations by stressing that they are interested in the way society views members of various categories. Other research has assessed people's beliefs about various subtypes within broader social categories (e.g., within the category of women, one could separately assess beliefs about housewives, businesswomen, old maids, etc.). Despite these improvements, there still exists some ambiguity in the interpretation of the resulting data.

These concerns have prompted researchers to devise more subtle and indirect ways of measuring stereotypes. These methods often involve subliminally presenting a picture or words related to the stereotype (activating or "priming" the stereotype) and then assessing whether the participants respond to another task in stereotype-consistent ways. For example, studies have primed the stereotype of African Americans and have then shown that participants take less time to recognize negative (compared to positive) words. Other studies have shown that they make more stereotypic completions of word fragments (e.g., _a_y = lazy). These and other indirect techniques have been used to assess people's thoughts about various groups without being subject to social desirability biases.

In summary, stereotypes may be measured in a number of ways, each of which has its strengths and limitations.

5. WHEN DO WE USE STEREOTYPES?

From categorization processes to automatic activation, stereotypes are used by everyone at some time or another. However, what determines whether or not we will use stereotypes when forming impressions of another person? Results of a number of studies on this topic have determined that stereotypes are most likely to influence impressions of another person when that person's behavior is ambiguous and therefore is open to stereotype-congruent interpretations. For example, seeing someone spend many hours studying for an exam, a perceiver's interpretation of that person's effort might be guided by racial stereotypes. If the student is Asian, then the extensive studying might be attributed to a desire to maintain a high grade point average. On the other hand, if the student is African American, then the perceiver might assume that the person is struggling to pass the class and therefore must study hard.

In contrast to this situation in which there are multiple possible interpretations of the behavior, there are other times when one's behavior is a direct indication of his or her underlying traits or dispositions. For example, a person who receives very high scores on the GRE exam would be seen as being intelligent, regardless of the person's race, gender, and so on.

The results of this research, then, have shown that stereotypes influence impressions of others when the behavior in question is ambiguous and open to stereotype-consistent interpretations. However, their influence on our construal of behavior is not the only consequence of using stereotypes.

6. CONSEQUENCES OF STEREOTYPING

6.1. Loss of Individual Information

Thinking of stereotypes as mental shortcuts seems not only efficient and useful but also fairly harmless.

However, holding stereotypic beliefs about group members can have a number of important, and often negative, consequences. On a cognitive level, stereotypes help us group similar others into meaningful categories. Although useful, the mere act of placing people into a category means that you overgeneralize and lose individuating information about them. In other words, the individual is perceived as having all of the characteristics stereotypic of the group and thereby loses anything that might differentiate the person as a unique individual.

At the group level, people tend to view members of the out-group as more similar to each other than members of the in-group (known as the out-group homogeneity effect). Perceiving homogeneity in a group is an important step in stereotyping because if all group members are seen as very similar, then it becomes easy to overgeneralize about them. Also, as already mentioned, overgeneralization is one of the hallmarks of stereotyping. In contrast, members of one's in-group retain their individuality and are seen as distinct from each other. A consequence of this in-group/out-group distinction is that people have a tendency to favor their own group (in-group bias) rather than the out-group. This in-group bias then becomes harmful if it results in prejudice toward out-group members or in favoritism toward in-group members based solely on their membership in the in-group.

6.2. Behavioral Implications

Once activated, stereotypes may influence both our behavior and our interpretation of another's actions. As mentioned previously, stereotypes act as expectancies that influence our construal of another's ambiguous behavior. In addition, they may unconsciously influence us to behave in ways that are consistent with the stereotype. For example, several studies have shown that priming the African American stereotype leads participants to behave in a more hostile way and to interpret another person's behavior as more aggressive. Subtly priming young people with the stereotype of the elderly can lead those young people to move more slowly. These effects become a very real problem when, even unconsciously, our beliefs about group members influence both our behavior toward them and, in turn, their subsequent behavior toward us.

6.3. Self-Fulfilling Prophesy

The process by which one's beliefs about a person lead that person to behave in ways that confirm one's original expectations is called the self-fulfilling prophesy. The importance of the self-fulfilling prophesy was highlighted in a 1968 study by Rosenthal and Jacobsen. They led teachers to believe that a randomly chosen sample of children in their classes were very intelligent and were about to blossom. The research showed that the teachers then treated those students in ways that helped the students perform well, thereby confirming the teachers' expectations. Thus, the teachers' experimentally induced expectations guided their own behavior, which in turn produced the expected outcome in the students' performance.

These effects have been replicated in a number of other situations (including interactions between males and females and between blacks and whites). In each case, people's stereotypic beliefs about a target person influenced that person's behavior, thereby confirming the people's original expectations. These studies stress the important role that stereotypes play in several aspects of the interpersonal context. They not only can influence perceptions of behavior and guide the perceiver's own behavior but also can influence the stereotyped person to behave in a stereotype-confirming manner.

7. STEREOTYPE CHANGE

With the knowledge that stereotypes are often based on incomplete overgeneralizations, it seems that one obvious way to change people's stereotypes is to put them into contact with the very people they are stereotyping. Specifically, by interacting with out-group members people would have ample opportunity to notice behavior that was inconsistent with the group stereotype, thereby resulting in stereotype change. This premise is exactly what Allport's 1954 contact hypothesis is based on. Although the idea is simple, there are certain factors that influence whether or not contact leads to stereotype change. For example, if a certain white individual holds negative stereotypes about Latinos, then, according to the contact hypothesis, those stereotypes should be reduced by having the individual interact with Latinos in a supportive, friendly environment. In addition, stereotype change is facilitated when social institutions support the intergroup contact and when the two groups are believed to share equal status in some domain. Clearly, not every interaction with members of the stereotyped group will lead to change. In fact, situations in which the contact involves competition, members of one or both groups are frustrated, and members of the

two groups hold different morals or values will not be conducive for reducing stereotypes.

In the multicultural society in which we live, people of different races seem to be interacting with increasing frequency, yet stereotypes persist. One possible explanation for this persistence is that people rationalize inconsistent behaviors by stereotyped group members. For example, if someone meets a successful Latino businessman, then this might go against his or her stereotypic expectation that Latinos are lazy and unintelligent. However, rather than changing their stereotypes, people may instead group people that do not "fit" the stereotype into their own separate category or subtype. By placing those who display stereotype-inconsistent behavior into subtypes, people maintain the status quo and resist changing their stereotypes of the group as a whole.

Although the picture painted previously seems grim, researchers have developed other promising cognitive interventions that have been shown to decrease the use of stereotypes in certain situations. For example, encouraging people to think carefully about others before passing judgment forces them to focus on individual-level attributes rather than category-level information. This approach encourages people to look beyond group memberships and instead to see people as unique individuals. Moreover, if the motivation and cognitive resources are available, people are able to control their use of stereotypes in many situations.

8. CONCLUSION

Stereotypes are multifaceted, complex phenomena that influence us all in one way or another. At their most basic level, stereotypes enable us to effortlessly separate people into distinct groups, thereby allowing us to make inferences about them. In the process of categorizing people into groups we may gain information, but we also overlook unique information about individuals. In addition to categorization, stereotypes play a number of dual roles. They may be positive or negative, and they often lead to in-group favoritism or out-group derogation. Despite the wide array of consequences from the stereotyping process, stereotypes can be controlled and changed if circumstances permit. Stereotypes are a fundamental human resource but may become harmful if categorization and subsequent overgeneralizations lead to prejudice and discrimination.

See Also the Following Articles

Groupthink ■ Intergroup Relations and Culture ■ Prosocial Behavior, Development of

Further Reading

Adorno, T. W., Frenkel-Brunswik, E., Levinson, D. J., & Sanford, R. N. (1950). *The authoritarian personality*. New York: Harper.

Allport, G. W. (1954). *The nature of prejudice*. New York: Addison-Wesley.

Bargh, J. A., Chen, M., & Burrows, L. (1996). Automaticity of social behavior: Direct effects of trait construct and stereotype activation on action. *Journal of Personality and Social Psychology*, *71*, 230–244.

Darley, J. M., & Gross, P. H. (1983). A hypothesis-confirming bias in labeling effects. *Journal of Personality and Social Psychology*, *44*, 20–33.

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, *56*, 5–18.

Eagly, A. H., & Steffen, V. J. (1984). Gender stereotypes stem from the distribution of women and men into social roles. *Journal of Personality and Social Psychology*, *46*, 735–754.

Gilbert, D. T., & Hixon, J. G. (1991). The trouble of thinking: Activation and application of stereotypic beliefs. *Journal of Personality and Social Psychology*, *60*, 509–517.

Hamilton, D. L., & Sherman, J. W. (1994). Stereotypes. In R. S. Wyer, Jr., & T. K. Srull (Eds.), *Handbook of social cognition* (2nd ed., Vol. 2, pp. 1–68). Hillsdale, NJ: Erlbaum.

Katz, D., & Braly, K. W. (1933). Racial stereotypes of 100 college students. *Journal of Abnormal Social Psychology*, *28*, 280–290.

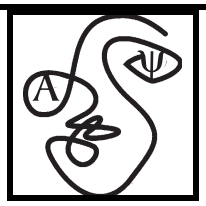
Lippman, W. (1922). *Public opinion*. New York: Harcourt Brace.

Rosenthal, R., & Jacobsen, L. (1968). *Pygmalion in the classroom*. New York: Holt, Rinehart, & Winston.

Snyder, M., Tanke, E. D., & Berscheid, E. (1977). Social perception and interpersonal behavior: On the self-fulfilling nature of social stereotypes. *Journal of Personality and Social Psychology*, *35*, 656–666.

Tajfel, H., & Turner, J. C. (1986). The social identity theory of intergroup behavior. In S. Worchel, & W. G. Austin (Eds.), *Psychology of intergroup relations* (pp. 7–24). Chicago: Nelson-Hall.

Word, C. O., Zanna, M. P., & Cooper, J. (1974). The nonverbal mediation of self-fulfilling prophesies in interracial interaction. *Journal of Experimental Social Psychology*, *10*, 109–120.



Stress

Charles J. Holahan and Jennifer D. Ragan
University of Texas, Austin, Texas, USA

Rudolf H. Moos
Department of Veterans Affairs and Stanford University, Menlo Park, California, USA

1. Introduction
2. Categories of Life Stressors
3. Health Effects
4. Stress Moderators
5. Conclusions
Further Reading

GLOSSARY

adaptive coping Cognitive and behavioral efforts to manage stressful conditions or associated emotional distress; adaptive coping efforts operate as protective factors that decrease the adverse effects of life stressors when they occur and that can also reduce the likelihood of stressor occurrence.

chronic stressors Stressors that endure over time (e.g., social role strains, everyday hassles).

hypothalamic–pituitary–adrenocortical axis Mechanism mediated by the pituitary gland, which during stressful experiences stimulates the adrenal cortex to release glucocorticoids that suppress key components of the immune system when chronically expressed.

negative life events Acute life change events that occur at discrete points in time and that are objectively rated as negative (e.g., death of family members, financial setbacks).

personal diatheses Long-standing biological or psychological vulnerabilities that interact with life stressors and increase initial health risk and that can also impede the mobilization of social resources and adaptive coping efforts; they may reflect both hereditary and developmental influences.

social resources Emotional support and assistance from family members and friends; these resources operate as protective factors that decrease the negative health effects of life stressors when they occur and that can also reduce the likelihood of stressor occurrence.

stress moderators Factors that increase or decrease the health effects of life stressors by functioning as either vulnerability factors or protective factors when stressors occur.

sympathetic arousal Mechanism mediated by the sympathetic nervous system, which during stressful experiences stimulates the release of catecholamines that mobilize the body to deal with immediate adaptive demands.

traumatic stressors Qualitatively more severe stressors than those indexed on standard life event or role strain scales (e.g., being exposed to war or a natural disaster, being the victim of a violent crime).

Life stressors are linked to both psychological and physical illnesses. Early stress studies focused on the direct negative effects of stressors on psychological and physical health, overlooking individual variability in response. However, later studies have emphasized factors that moderate the potential health effects of life stressors by functioning as either vulnerability factors or protective factors when stressors occur.

1. INTRODUCTION

The stress process is initiated by environmental factors (stressors) that place adaptive demands on individuals. Contemporary understanding of the physiology of stress

©2004 Elsevier Inc.
All rights reserved.

owes much to the groundbreaking research of Cannon on the role of the sympathetic nervous system in mobilizing the body for adaptive challenge and of Selye on the role of corticosteroids in producing the adverse biological effects of chronic stress. Stress research during the 1960s and 1970s extended early findings based on laboratory animals to humans, presuming a straightforward link between life change and pathology that overlooked individual variability in response. In fact, life stressors are linked to both psychological and physical illnesses. However, despite the consistency of stressor effects, early stress research was characterized by persistent anomalies. The amount of variance predicted in distress was typically less than 10%, and individuals showed highly variable reactions to stressors. Stressors made some individuals ill, others stayed healthy, and some grew stronger after managing challenges successfully.

Researchers recognized that these apparent anomalies reflected limitations in conceptualizing and measuring stressors and stress reactions; however, only slowly did they come to understand that these anomalies also represented important findings in their own right. The relatively poor empirical predictions of early stress–illness studies led researchers to focus increasingly on factors that moderate the health effects of life stressors. Contemporary research assumes that the human stress response is inherently complex, reflecting a dynamic interplay among stressors and vulnerability and protective factors in the individual and in the social environment.

2. CATEGORIES OF LIFE STRESSORS

Stress studies investigate (a) life change events, (b) major life crises and traumas, and (c) chronic stressors (Table I). It should be noted that although life change events have an acute onset, they often initiate chronic stressors such as persisting economic and family problems. In fact, these enduring secondary stressors may account for many of the adverse effects of life change events.

2.1. Life Change Events

Early stress studies focused on acute life change events that occur at discrete points in time such as deaths of family members and financial setbacks. A common measurement instrument was Holmes and Rahe's Social Readjustment Rating Scale (SRRS). The SRRS is a self-report questionnaire that uses objectively weighted life change units to index the relative level of adaptive demand associated with a particular life event, regardless of whether the event is positive or negative. These weighted life change units are summed to reflect cumulative life change during a given period of time, typically the previous 12 months. Later research revealed that the adverse health effects of life change are linked to negative life events rather than positive events. Later studies also showed that weighted life change units contributed little beyond an unweighted sum of events experienced in predicting illness. Based on these findings, most contemporary research on life change events indexes the sum of negative events experienced during a given period of time.

TABLE I
Examples of Life Stressors by Measurement Category

Life change events	*Traumatic stressors*	*Chronic stressors*
Death of spouse	War exposure	Family role strains
Death of friend	Victim of violent crime	Work role strains
Divorce	Natural disaster	Discrimination
Fired at work	Technological disaster	Poverty
Financial setback	Serious medical illness	Caregiving

2.2. Major Life Crises and Traumas

Stress researchers have also studied the adverse health effects of traumatic stressors that are qualitatively more severe than the stressors indexed on standard life event or role strain scales. Initially, traumatic stressors were defined as personal experience of life-threatening stressors such as being exposed to war and being the victim of a violent crime. Also included were being exposed to natural disasters (e.g., tornados, earthquakes) and being diagnosed with life-threatening medical illnesses (e.g., heart disease, cancer). Over time, the definition of trauma has been broadened to encompass witnessing events that place others at risk of death or serious injury or witnessing individuals being badly injured or killed by such events. Interest in traumatic stressors is evidenced in a growing body of research on posttraumatic stress disorder (PTSD).

2.3. Chronic Stressors

A major limitation of early stress research was its narrow preoccupation with laboratory studies of acute stressors.

Eventually, recognizing that many life stressors are ongoing, investigators began to examine chronic stressors that endure over time such as social role strains and everyday hassles. Researchers studying social role strains often measure chronic family or work difficulties. Some stress studies have focused on additional burdens that often are associated with family life such as caregiving for chronically ill family members. For example, increasing interest is being devoted to caregivers of aging family members suffering from Alzheimer's disease. Typical measurement instruments include Moos's Work Environment Scale and Family Environment Scale, which measure the social climate at work (e.g., work pressure) and in the family (e.g., family conflict), respectively. Moreover, there has been increasing research interest in role strains associated with the interface between work and family lives. Investigators have also studied minor life stressors that are experienced chronically such as everyday hassles with troublesome neighbors and traffic problems. Increasingly, stress studies are examining the health consequences of ongoing "macrostressors" such as discrimination, neighborhood violence, poverty, and terrorism.

3. HEALTH EFFECTS

Life stressors inevitably produce some emotional strain and physical tension. Stress researchers have posed the question of whether life stressors are also associated with more significant illnesses. In fact, life stressors are linked to a number of psychological and physical illnesses. Psychological and physical illnesses are often linked reciprocally, with each category of illness exacerbating the other. Moreover, psychological and physical illnesses often function as life stressors themselves, initiating a new cycle in the stress process.

3.1. Psychological Illness

Life stressors are associated with psychological stress reactions that involve depression and anxiety. For example, life stressors are linked both to the onset of depressive disorders and to relapse among individuals recovering from depressive disorders. Interpersonal problems and losses are especially likely to be associated with depressive reactions. Life stressors can also precipitate both onset and relapse of anxiety disorders such as generalized anxiety disorder, panic disorder, agoraphobia, and obsessive–compulsive disorder, and they can play a role in the development and progression of alcohol and drug abuse. Moreover, life stressors can trigger schizophrenic episodes among individuals who are vulnerable to this disorder. Trauma exposure produces a recognized pattern of PTSD symptoms, including reexperiencing the trauma psychologically through flashbacks and nightmares, emotional numbing, and experiencing heightened arousal and vigilance.

Exclusive of traumatic events, chronic stressors are more strongly linked to psychological distress than are acute events. Although chronic stressors generally are less severe than acute life events, their effects often last longer and are more pervasive (Fig. 1). Moreover, an event is more likely to have an adverse psychological outcome when it threatens or disrupts a domain in which a person has central commitments.

Psychological reactions themselves can exacerbate the stress process in two important ways. First, psychologically distressed individuals are more likely to perceive benign situations as threatening, and these perceptions of threat can trigger additional stress reactions. Second, persons who are psychologically distressed often create social conditions in their lives, such as conflictual family or work relationships, that are likely to produce new life stressors.

3.2. Physical Illness

Stressor exposure also initiates a characteristic biological response that is associated with the onset or exacerbation of a wide spectrum of physical illnesses. The biological stress response involves interconnections among the nervous, endocrine, and immune systems.

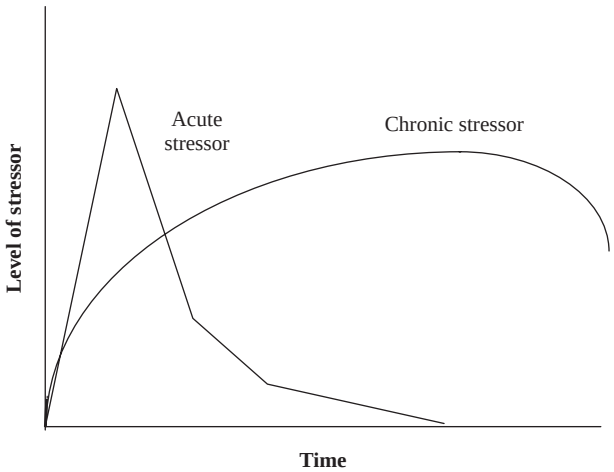


FIGURE 1 The time course of acute stressors versus chronic stressors. Adapted from Wheaton (1994).

The two most heavily studied stress-related biological mechanisms have been sympathetic arousal and activation of the hypothalamic–pituitary–adrenocortical (HPA) axis. Both mechanisms involve initial central nervous system input from the hypothalamus at the base of the brain, and both mechanisms operate through the adrenal glands located above the kidneys (Fig. 2). Short-term activation of these biological mechanisms is adaptive, mobilizing energy and enhancing alertness to respond to adaptive demands. However, chronic stimulation of these mechanisms can lead to (a) hyperarousal and subsequent wear and tear on body systems and (b) suppression of key components of the immune system.

Sympathetic arousal is mediated by the sympathetic nervous system, which during stressful experiences releases norepinephrine at multiple sites throughout the body and stimulates the adrenal medulla to release epinephrine (adrenaline). These catecholamines mobilize the body to deal with immediate adaptive demands. However, chronic sympathetic arousal is associated with tension-related complaints such as headache. In fact, stress is the triggering factor reported most often by migraine and tension-type headache patients. Sustained or frequent intense sympathetic arousal is also linked to wear and tear on arteries and coronary vessels and to shear stress associated with sharply increasing catecholamines. Chronic sympathetic arousal also increases blood clotting through coronary vasoconstriction, increased circulating lipids, and increased platelet aggregation.

Hemodynamic and biochemical changes, in turn, may lead to health-related problems such as atherosclerosis, hypertension, and coronary artery disease. For example, occupational stressors have been linked to elevated blood pressure in studies of activity at work. Moreover, stressful work conditions, particularly the combination of high work demands and low job control, have been associated with coronary heart disease in population studies of workers in Europe and the United States.

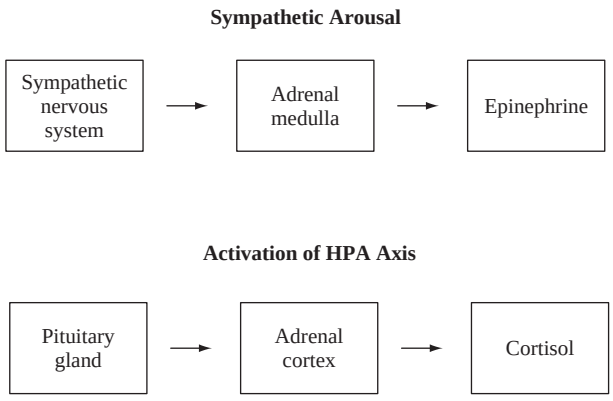


FIGURE 2 Two stress-related biological mechanisms mediating between central nervous system input and health outcomes. HPA, hypothalamic–pituitary–adrenocortical axis.

Activation of the HPA axis is mediated by the pituitary gland, which during stressful experiences stimulates the adrenal cortex to release glucocorticoids (primarily cortisol in humans). Chronic expression of glucocorticoids and excessively high levels of glucocorticoids have been associated with suppression of key components of the immune system in human in vitro studies (i.e., examining tissue or blood samples outside the body) and in experimental studies using animal models. For example, glucocorticoids are linked to reduced T cell proliferation and lower natural killer cell cytotoxicity. In addition, glucocorticoids can impede the maturation of developing lymphocytes and can destroy mature lymphocytes. These changes inhibit the ability of the cellular arm of the immune system to attack target pathogens directly. Glucocorticoids are also linked to decreased production of antibodies and certain components of complement. These changes inhibit the ability of the humoral arm of the immune system to target pathogens for destruction. It is unclear, however, whether the levels of immune system downregulation observed in the context of stressful experiences are sufficient to increase vulnerability to disease.

In studies examining clinical illness, there has been evidence of a link between life stressors and the onset and progression of infectious diseases such as the common cold and influenza. Life stressors are also related to latent viral activity in the herpes viruses, resulting in complaints such as mononucleosis and cold sores. In addition, life stressors may be associated with the onset and exacerbation of some autoimmune diseases, such as rheumatoid arthritis, and with poor diabetic control and progression of diabetes. Experimental research with animals has demonstrated that stress can increase the development of artificially induced tumors; however, there is no clear evidence that stress can accelerate the growth of common human tumors.

Historically, stress research has emphasized the link between sympathetic arousal and cardiovascular response as well as that between the HPA axis and immune system response. However, there is increasing evidence that sympathetic arousal also affects the immune system and that the HPA axis also affects the cardiovascular system. In addition to these direct

biological pathways between stressors and illness, life stressors can relate to physical illness indirectly through health risk behaviors. Life stressors are associated with common health risk behaviors such as increased cigarette smoking, alcohol abuse, poor dietary and exercise habits, disturbed sleep, and reduced adherence to medical regimens. In turn, each of these health risk behaviors plays a role in increasing the risk of physical illness.

4. STRESS MODERATORS

Individuals have highly variable reactions to stressors, with some individuals showing little or no health effects from stressor exposure and others showing profound reactions. From an initial emphasis on a straightforward link between life stressors and illness, contemporary stress studies have evolved to placing increasing emphasis on stress moderators. Moderating factors increase or decrease the health effects of life stressors by functioning as either vulnerability factors or protective factors when stressors occur.

More recent stress research has been guided by two models that emphasize the key role of stress moderation: the diathesis–stress model and the stress resistance model. Figure 3 shows an integrative framework of key components of these two models. The diathesis–stress model focuses on long-standing vulnerability factors (i.e., personal diatheses) and emphasizes the first stage of the stress process where personal vulnerabilities interact with life stressors and increase initial health risk. The stress resistance model focuses on protective factors (i.e., social resources and adaptive coping efforts) and emphasizes the second stage of the stress process where protective factors mobilized after stressor exposure modify initial health risk and determine the final health outcome. The figure is simplified in that arrows are shown only for primary relationships. Personal vulnerabilities can affect the second stage of the stress process by impeding the mobilization of social resources and adaptive coping efforts, whereas protective factors can affect the first stage of the stress process by reducing the likelihood of stressor occurrence.

4.1. Personal Diatheses

Personal diatheses are long-standing biological or psychological vulnerabilities that interact with life stressors and increase initial health risk (Fig. 4). Personal diatheses can also impede the mobilization of social resources and adaptive coping efforts. Thus, adaptation can be understood in terms of the match between the level of life stressors and individuals' personal vulnerabilities. For example, individuals with many personal vulnerabilities typically adapt poorly except in settings with exceptionally low life stressors. Conversely, individuals with few personal vulnerabilities typically adapt well except in settings with exceptionally high life stressors. Individuals with an intermediate level of personal vulnerabilities typically adapt poorly in circumstances that tax their particular areas of fragility.

Personal diatheses reflect both hereditary and developmental influences and can affect both psychological and physical health outcomes under stress. For example, excessive cardiovascular reactivity (a biological

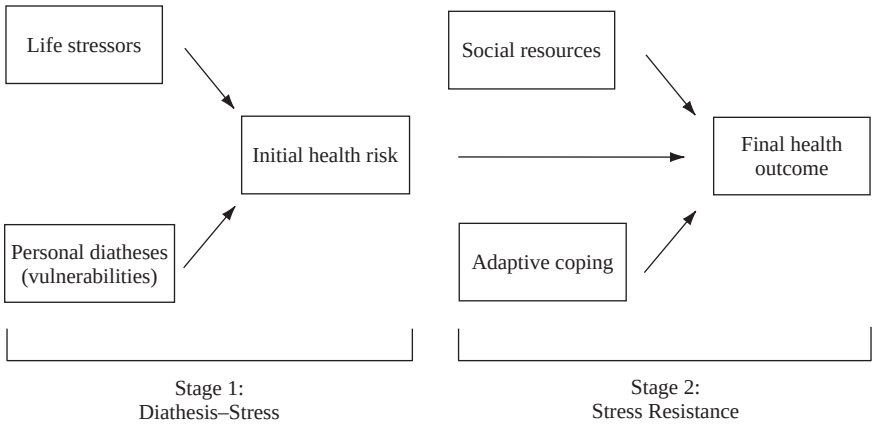


FIGURE 3 Integrative framework of key components of the diathesis–stress and stress resistance models. For simplicity, arrows are shown only for primary relationships.

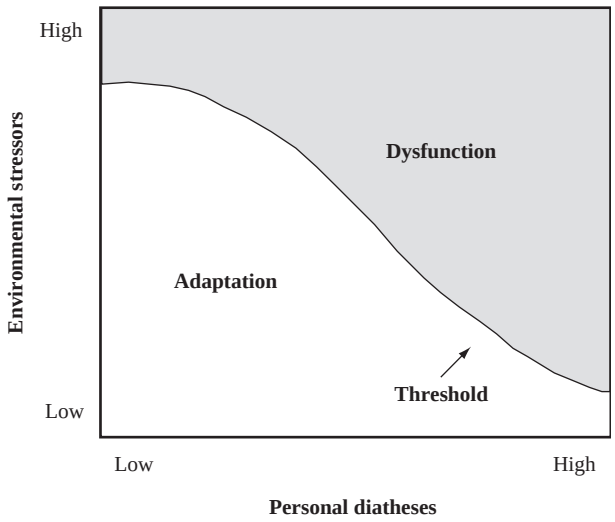


FIGURE 4 A matching model predicting psychological adjustment as a function of both environmental demands and personal vulnerability. Adapted from Zubin, J., & Spring, B. (1977). Vulnerability: A new view of schizophrenia. *Journal of Abnormal Psychology, 86,* 103–126. Copyright 1977 by the American Psychological Association. Reprinted by permission.

vulnerability) and hostility/cynicism (a psychological vulnerability) increase the risk of coronary artery disease under stress. Correspondingly, abnormalities in neurotransmitter systems, such as the serotonergic system in the brain (a biological vulnerability) and a proclivity to attribute failure to personal deficits (a psychological vulnerability), increase the risk of depressive disorders under stress. Although the diathesis–stress model emphasizes vulnerabilities, individuals also bring long-standing personal strengths to the stress process. Especially important are personality characteristics, such as dispositional optimism and self-efficacy, that enhance perceptions of personal control during recovery from stressors and that can also reduce the likelihood of stressor occurrence. The study of protective personality strengths is encompassed under the stress resistance model along with social resources and adaptive coping.

4.2. Social Resources

Social resources include emotional support as well as guidance and assistance from family members and friends. Social resources operate as protective factors that decrease the negative health effects of life stressors when they occur and that can also reduce the likelihood of stressor occurrence (Fig. 5). Social resources can reduce both adverse psychological health outcomes, such as depression after bereavement, and adverse physical health outcomes, such as job-related hypertension. Some aspects of social support, such as social companionship, may be related to better health and well-being regardless of the level of stressors experienced. Other aspects of social support, such as informational support, may function as "stress buffers" that have stronger positive effects on health when individuals are experiencing a high level of stressors. Although the stress resistance model emphasizes resilience, social relationships can also present health risks. In fact, negative aspects of relationships, such as criticism or disapproval from family members or friends, are as strongly related to poorer health under stress as positive aspects of relationships are related to better health under stress.

Recently, stress studies have examined how life stressors can deplete social resources. In part, such depletion may be tangible in that negative life events, such as the deaths of family members or friends, entail reductions in social resources. However, resource depletion also may have a cognitive component in that perceived support is vulnerable to deterioration after life crises. Moreover, resource loss demonstrates another way in which life stressors are linked to adverse psychological outcomes. For example, individuals recovering from a natural disaster might feel that the support they received from friends was less than what they had expected. This perception of resource loss mediates part of the link between crisis exposure and emotional distress.

4.3. Adaptive Coping

Adaptive coping includes cognitive and behavioral efforts to manage stressful conditions or associated emotional distress. Like social resources, adaptive coping operates as a protective factor that decreases the adverse effects of life stressors when they occur and that can also reduce the likelihood of stressor occurrence. People who rely more on active coping strategies, such as problem solving and information seeking, tend to adapt better to life stressors. Although the stress resistance model emphasizes resilience, coping styles may also involve vulnerabilities. For example, avoidance coping strategies, such as denial and wishful thinking, are generally associated with increased psychological distress. In addition, coping styles can affect physical health indirectly through these psychological reactions. Moreover, coping styles are linked to physical health indirectly through health risk behaviors

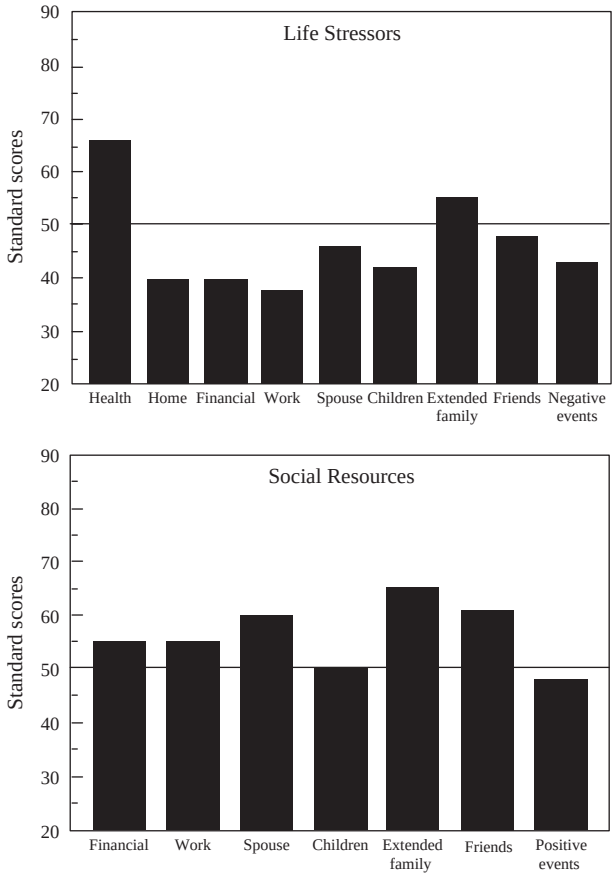


FIGURE 5 A Life Stressors and Social Resources profile for a woman with rheumatoid arthritis. Moos and Moos's Life Stressors and Social Resources Inventory (LISRES) provides a picture of the overall balance between an individual's life stressors and social resources across representative life domains. The inventory consists of nine indexes of life stressors and seven indexes of social resources. The figure shows an illustrative LISRES profile for a 66-year-old woman with rheumatoid arthritis (scores are standardized, with a mean of 50 and a standard deviation of 10). Although she reported well above-average physical health stressors, she had below-average stressors in six of the other seven domains. Moreover, she also had above-average social resources in her work setting and in her relationships with her spouse, extended family, and friends. At an 18-month follow-up, this favorable balance of moderate stressors and high resources enabled her to manage quite well. She reported high self-confidence and below-average depression. Adapted from Moos, R. H., & Moos, B. S. (1994). *Life Stressors and Social Resources Inventory: Adult form manual*. Lutz, FL: Psychological Assessment Resources. Reproduced by special permission of the publisher, Psychological Assessment Resources Inc., 16204 North Florida Avenue, Lutz, FL 33549, from the LISRES-A by Rudolf Moos, Ph.D., and Bernice Moos, copyright 1994 by PAR Inc. Further reproduction is prohibited without permission of PAR Inc.

such as cigarette smoking, alcohol use, and lack of adherence to medical regimens.

Based on these observations, most conceptualizations of coping emphasize the distinction between strategies that are oriented toward approaching and confronting the problem and strategies that are oriented toward reducing tension by avoiding dealing directly with the problem (Table II). Lazarus and Folkman proposed a related typology of coping processes that distinguishes between problem-focused strategies (i.e., efforts to modify the source of the stress) and emotion-focused strategies (i.e., attempts to regulate the emotional distress caused by a stressor). Emotion-focused strategies can be oriented toward dealing with a stressor; however, such coping is most often directed toward avoiding dealing with the source

TABLE II
Basic Coping Categories with Associated Coping Subtypes from Moos's Coping Responses Inventory

Basic coping category	*Coping subtypes*
Cognitive approach coping	Logical analysis (e.g., "Did you think of different ways in which to deal with the problem?")
	Positive reappraisal (e.g., "Did you think about how you were much better off than other people with similar problems?")
Behavioral approach coping	Seeking guidance and support (e.g., "Did you talk with a friend about the problem?")
	Taking problem-solving action (e.g., "Did you make a plan of action and follow it?")
Cognitive avoidance coping	Cognitive avoidance (e.g., "Did you try to forget the whole thing?")
	Resigned acceptance (e.g., "Did you lose hope that things would ever be the same?")
Behavioral avoidance coping	Seeking alternative rewards (e.g., "Did you get involved in new activities?")
	Emotional discharge (e.g., "Did you yell or shout to let off steam?")

Source. From Moos, R. H. (1993). *Coping Responses Inventory: Adult form manual* (3rd ed.). Lutz, FL: Psychological Assessment Resources. Adapted and reproduced by special permission of the publisher, Psychological Assessment Resources Inc., 16204 North Florida Avenue, Lutz, FL 33549, from the Coping Responses Inventory by Rudolf Moos, Ph.D., copyright 1993 by PAR Inc. Further reproduction is prohibited without permission of PAR Inc.

Note. Sample coping responses are in parentheses.

of stress. Although one can draw overall conclusions about the relative efficacy of approach and avoidance coping strategies, such generalizations necessarily oversimplify the process of adaptation. Individuals adapt best when their coping efforts match situational demands. Individuals who are flexible in their choice of coping should show better adaptation than do persons who have more restricted or rigid coping repertoires. For example, avoidance coping can be adaptive in the short term with time-limited stressors and with stressors that are not controllable.

Stress studies have begun to develop an integrated understanding of how social resources and adaptive coping are related to one another. Social resources are linked to health under high stressors in part because they encourage more adaptive coping strategies. For example, social resources can enhance coping efforts by bolstering feelings of self-confidence as well as by providing informational guidance that aids in appraising threat accurately and in planning effective responses.

5. CONCLUSIONS

Considerable evidence documents that life stressors are associated with both psychological and physical illnesses. However, the overall association between life stressors and illness is only small to moderate. Moreover, health effects are strongest with chronic stressors and severe stressors, which often precipitate associated ongoing demands. Even more important, the profound individual differences in stressor effects highlight the complexity of the stress process. At initial stressor exposure, potential health risk is moderated by long-standing biological and psychological vulnerabilities. Further moderation occurs after stressor exposure as social resources and individual coping strengths are brought into play to manage adaptive demands.

Finally, despite the pain of stressful experiences, some individuals are able to grow stronger under stress. For example, physiological resilience can be strengthened under stress. Physiological strengthening is most likely when stressors are acute rather than chronic and when periods of relaxation provide restorative intervals between stressor exposures. Similarly, life crises can be "constructive confrontations" that provide an opportunity for psychological growth. Individuals can emerge from a crises with greater self-confidence, new coping skills, closer relationships with family members and friends, and a richer appreciation of life.

See Also the Following Articles

Coping ■ Posttraumatic Disorders

Further Reading

Baum, A., & Posluszny, D. M. (1999). Health psychology: Mapping biobehavioral contributions to health and illness. *Annual Review of Psychology*, *50*, 137–163.

Cohen, S., & Herbert, T. B. (1996). Health psychology: Psychological factors and physical disease from the

perspective of human psychoneuroimmunology. *Annual Review of Psychology*, *47*, 113–142.

Coyne, J. C., & Downey, G. (1991). Social factors and psychopathology: Stress, social support, and coping processes. *Annual Review of Psychology*, *42*, 401–425.

Kessler, R. C. (1997). The effects of stressful life events on depression. *Annual Review of Psychology*, *48*, 191–214.

Krantz, D. S., & McCeney, M. K. (2002). Effects of psychological and social factors on organic disease: A critical assessment of research on coronary heart disease. *Annual Review of Psychology*, *53*, 341–369.

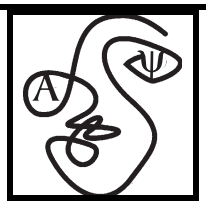
Lazarus, R. S. (1993). Coping theory and research: Past, present, and future. *Psychosomatic Medicine*, *55*, 234–247.

Moos, R. (1994). *The social climate scales: An overview*. Palo Alto, CA: Consulting Psychologists Press.

Sapolsky, R. M. (2002). Endocrinology of the stress response. In J. B. Becker, S. M. Breedlove, D. Crews, & M. M. McCarthy (Eds.), *Behavioral endocrinology* (2nd ed., pp. 409–450). Cambridge, MA: MIT Press.

Selye, H. (1993). History of the stress concept. In L. Goldberger, & S. Breznitz (Eds.), *Handbook of stress: Theoretical and clinical aspects* (2nd ed., pp. 7–17). New York: Free Press.

Wheaton, B. (1994). Sampling the stress universe. In W. R. Avison, & I. H. Gotlib (Eds.), *Stress and mental health: Contemporary issues and prospects for the future* (pp. 77–114). New York: Plenum.



Structure of Interests

Christy Hofsess and Terence J. G. Tracey
Arizona State University, Tempe, Arizona, USA

1. Evolution of the Structure of Interests
2. Extensions of the Circular Model: Spherical Structure
3. Relation of the RIASEC Types to Personality and Ability Measures
4. Current Research
Further Reading

GLOSSARY

circumplex Model in which the focus is on a pattern of relations in a matrix that forms a circular structure. Variables within a specific domain (e.g., vocational interest domain) can be understood and described by their placement on the circle; those more proximate are more highly related than those more distal.

factor list Model in which the focus is on the number and nature of independent factors that exist in a particular domain (e.g., vocational interest domain).

hierarchical model Model in which the focus is on the discrete sets of clusters within a domain (e.g., vocational interest domain).

interest inventory An assessment of one's interests for use in counseling and guidance.

interests Activities and occupations that an individual finds attractive.

structure How something is constructed, organized, or interrelated.

The structure of interests is a well-researched area within vocational psychology that focuses on how individuals think about and conceptualize their interests. All theories are based on how interests are interrelated and aim to create a systematic framework for understanding the organization of interests, thus guiding interest inventory development and career counseling practice. Since the early 1950s, researchers have primarily developed theories categorizing and classifying the many types of interests and occupations individuals pursue into multiple factors or themes, and these are useful as organizing principles in inventory development and presentation. Interest inventories help individuals understand how their curiosities and involvement in different activities influence and shape career decisions. The focus on the structure of interests thus centers on gaining a better understanding of how individuals think about their interests and presenting more helpful information to the test taker.

1. EVOLUTION OF THE STRUCTURE OF INTERESTS

Much of the research on interests has been of the factor list variety as exemplified by the early work of Guilford in 1954 and others. Interest items were factor analyzed with the goal of yielding a parsimonious set of factors that characterize interests. This approach is useful but often results in numerous factors, making easy application somewhat difficult. As an alternative to the factor list approach, Roe introduced the first circular theory of

©2004 Elsevier Inc.
All rights reserved.

the structure of interests in 1956 by proposing a circular model to explain how interest types and occupations are related to each other. Her theory classified occupations into eight groups: service, business contact, organization, technology, outdoor, science, general culture, and arts and entertainment. These groups, referred to as fields, represent a common shared activity. Her theory arranged the fields in a circular manner so that similar fields are in close proximity. In addition to the circular arrangement of interests, she also introduced the concept of level of responsibility (analogous to complexity of tasks). Subsequent work by Tracey and Rounds demonstrated that the circular arrangement proposed by Roe does not have empirical support. However, Roe's model served as a stimulus to the circular model proposed by John Holland.

1.1. Holland's Model of Vocational Types

John Holland's theory of vocational interest types evolved into the dominant, standard organizing framework for interest assessment used by career counseling professionals. His theory classifies individuals and work environments according to six types: realistic (R), investigative (I), artistic (A), social (S), enterprising (E), and conventional (C), commonly referred to collectively as RIASEC. Although RIASEC types are primarily personality types, they typically are conceptualized and measured in terms of vocational interests. Holland posited these six types as existing in a hexagonal (or circumplex) structure with equal distance between the types, and proximity of the types representing the degree of similarity. His model is represented in Fig. 1. Individuals are categorized into the six types according to their personality dispositions. Similarly, work environments are characterized according to settings that evoke, develop, and reinforce the six personality interest types. The congruence of interest types and work environments is hypothesized to be related to satisfaction and productivity. The simplicity, elegance, and applicability of Holland's typology make it the most widespread theory of the structure of interests.

The prominence of Holland's theory has led to the incorporation of RIASEC scales into every major interest assessment instrument, such as the Strong Interest Inventory (SII), Self-Directed Search (SDS), American College Testing Program Interest Inventory (UNIACT), Kuder Occupational Interest Survey, Vocational Preference Inventory (VPI), and the Career Assessment Inventory.

1.2. Alternative Approaches

Although Holland's model remains the most influential model within vocational psychology, two other notable structural models diverge from his hexagonal structure: the hierarchical model of Gati and the dimensional model of Prediger. Using the RIASEC types, Gati proposed a hierarchical model of the structure of vocational interests as a more accurate representation of the structure of the six types than the circular/hexagonal representation. His hierarchical model is also depicted in Fig. 1. Tracey and Rounds compared Gati's hierarchical model to Holland's circular model in a meta-structural analysis using 104 different RIASEC correlation matrices with a total sample size of 47,268 individuals. They found that Holland's circular model was better supported than was Gati's model, especially for U.S. samples. Subsequent work by these researchers on cross-cultural and ethnic minority U.S. samples provides some support for the viability of Gati's model in these contexts.

Prediger sought to explicitly label the underlying dimensions of the RIASEC circular model. His model, although an alternative approach, complements Holland's theory by demonstrating how the RIASEC circular structure contains two key underlying work task dimensions, people–things and data–ideas. The people–things dimension differentiates the social type from the realistic type, and the data–ideas dimension separates enterprising and conventional types from artistic and investigative types. His model is summarized in Fig. 1. Prediger posited that a focus on the two underlying dimensions rather than the six types provides a more parsimonious representation of an individual's interests and the work environment.

In a different meta-structural analysis, Rounds and Tracey compared the circular model to Prediger's dimensional approach. They confirmed that underlying Holland's circular structure are two dimensions, supporting Prediger's people–things/data–ideas model as complementing, not competing with, Holland's theory. The focus of Holland's structure is less on the underlying dimensions and more on the circular structure, yielding information about the relations among the interest types not revealed in the dimensional approach. Therefore, although the two dimensions are valid, they are viewed as supplementary because they are not as important as the circular structure.

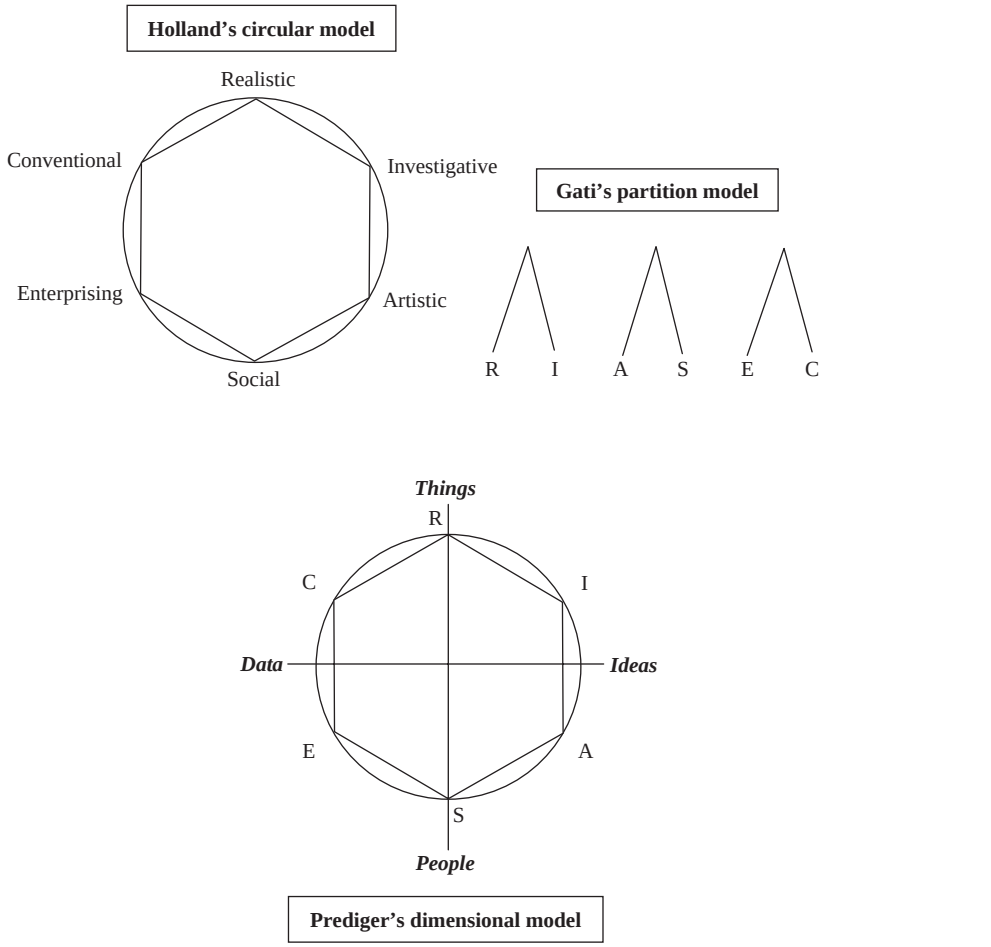


FIGURE 1 Various structural representations of Holland's RIASEC types (Holland circumplex, Gati partition, and Prediger dimensional). From Tracey, T. J. G. (1993). Evaluating Holland's and Gati's vocational interest models: A structural meta-analysis. *Psychological Bulletin, 113,* 229–246. Copyright © 1993 by the American Psychological Association. Adapted with permission.

1.3. Generalizability of the Circular Structure

Most RIASEC interest score comparisons across gender, culture, and age focus on mean differences in scores and do not take into consideration variations that may exist in underlying interest structure. In an investigation of the structural invariance of the circular order model, Tracey and Rounds found no differences in gender or age (at least with respect to samples aged 14 years or older). The circular structure was also invariant across the four major interest inventories (VPI, SDS, UNIACT, and SII). As noted previously, the circular model was the best representation of U.S. samples, whereas Gati's model appeared to be a better representation of non-U.S. samples.

2. EXTENSIONS OF THE CIRCULAR MODEL: SPHERICAL STRUCTURE

In two studies on the structure of interests, Tracey and Rounds demonstrated (i) the arbitrary nature of the RIASEC types and (ii) the presence of prestige in interest data. Interest items were found to be uniformly distributed around the interest circle, not centered on the six nodes of the RIASEC types. This implied that the

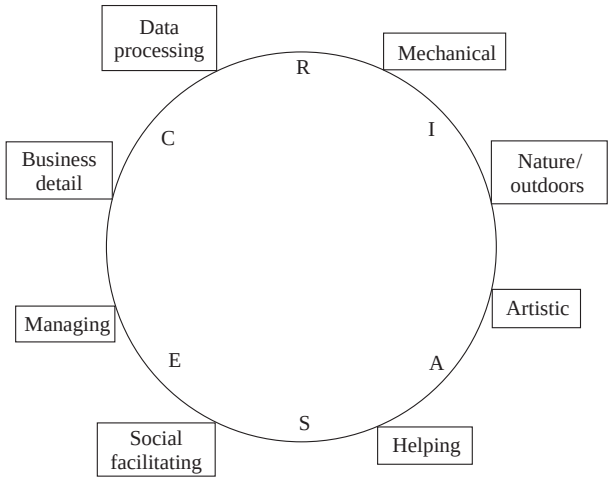


FIGURE 2 Interest circle represented by Holland's six RIASEC types and Tracey's eight types.

RIASEC hexagon can be better represented by a circle composed of any number of types and the adoption of six is arbitrary. To examine this, Tracey and Rounds created an eight type model and found that it fit the data as well as or better than the six type RIASEC model. A version of their eight type model is presented in Fig. 2.

Prestige—also referred to as status, occupational level, level of training, and level of difficulty or responsibility—is a salient variable in vocational assessment, but it is commonly unaccounted for in interest inventories. Adding prestige to the RIASEC interest circle, Tracey and Rounds proposed a spherical structure wherein a prestige dimension exists orthogonal to Prediger's people–things and data–ideas dimensions. Tracey's Personal Globe Inventory is an inventory developed to measure the spherical model and a representation of the structure is depicted in Fig. 3.

The model is best conceptualized as an occupational interest globe, with prestige represented along the north pole–south pole axis and serving to differentiate interest areas. For example, as one moves up the globe toward high prestige interest areas (e.g., social science, science, and financial analysis) or down the globe toward low prestige interest areas (e.g., personal service, basic services, and construction/repair) the salience of the people–things and data–ideas dimensions diminishes. Individuals with moderate levels of prestige preference would be characterized by standard RIASEC measures and appear near the equator on the occupational interest globe.

3. RELATION OF THE RIASEC TYPES TO PERSONALITY AND ABILITY MEASURES

Many theories suggest that vocational interests are an extension of personality and that interests are related to abilities. Holland used trait descriptive adjectives to describe the RIASEC interest types and claimed that they were essentially personality types. The most well-known theory of personality trait structure, the Big Five Model, is a factor list taxonomy for organizing personality variables. Correlations between the RIASEC and the Big Five are small to moderate and also mixed; enterprising and social types are correlated with extroversion, and artistic types are correlated with openness to experience. Research has generally ignored the circular structure of the RIASEC, with the exception of one study. One of the Big Five traits, extroversion, fit the RIASEC circular structure, separating enterprising and social types from realistic and investigative. Integrative work by Ackerman and Heggestad has focused on representing the overlap among the RIASEC interests, personality traits, and ability. They present initial models of the relation among each of these domains, and this research should serve as stimuli for subsequent research on the intersection of these three important domains.

4. CURRENT RESEARCH

Most research on the structure of interests focuses on adolescents and adults. Very little is known about the structure of interests in children or the development of the structure of interests. Initial studies demonstrate that the structure of interests in elementary school children does not mirror the circular structure found in those 14 years old or older. Work in the area of the development of interests and their structure is needed as an aid to design educational programs appropriate to the elementary school age group. Since the development of interests in children is inherently changeable and unique, research focusing on this process should examine differences in how children think about their likes and dislikes and how the development of interests changes over time.

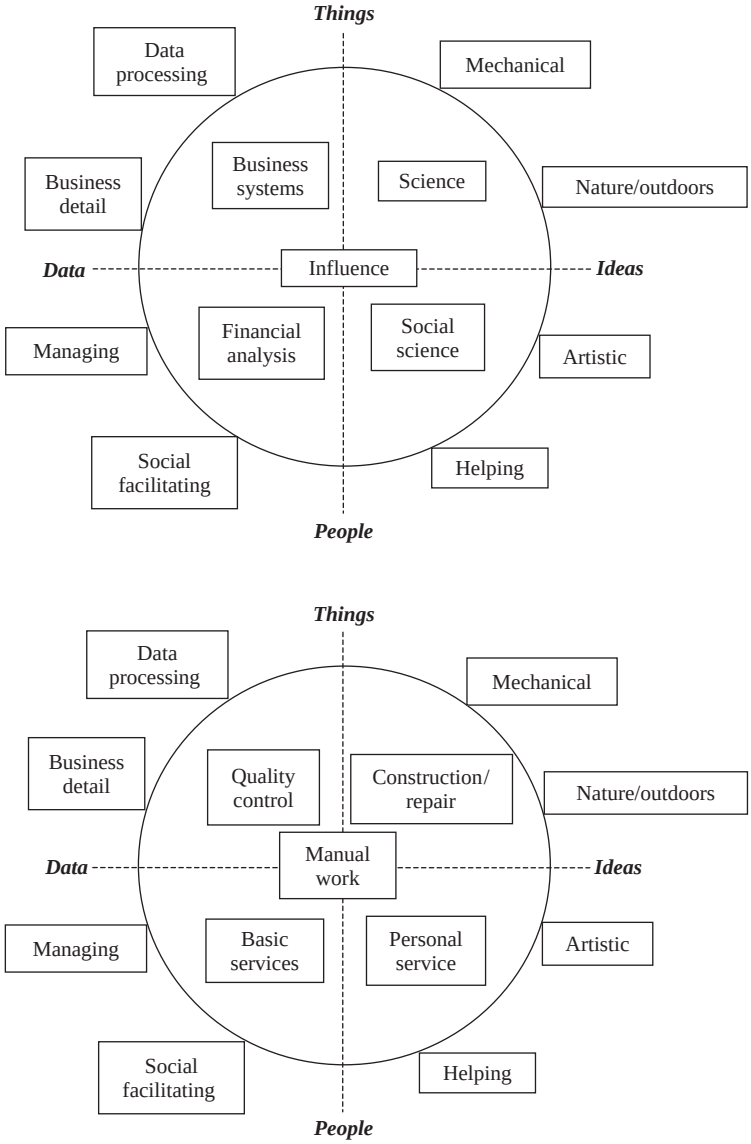


FIGURE 3 Spherical model of interests (as exemplified by the Personal Globe Inventory). (Top) Upper half (higher prestige); (bottom) lower half (lower prestige).

See Also the Following Articles

Career Counseling ■ Holland's Theory (Vocational Personality Types) ■ Indecision, Vocational ■ Job Search ■ Traits ■ Vocational Interests ■ Vocational Psychology, Overview

Further Reading

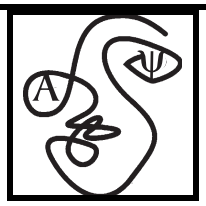
Gati, I. (1991). The structure of vocational interests. *Psychological Bulletin*, *109*, 309–324.

Holland, J. L. (1997). *Making vocational choices* (3rd ed.). Odessa, FL: Psychological Assessment Resources.

Prediger, D. J., & Vansickle, T. R. (1992). Locating occupations on Holland's hexagon: Beyond RIASEC. *Journal of Vocational Behavior*, *40*, 111–128.

Tracey, T. J. G. (2002). Personal Globe Inventory: Measurement of the spherical model of interests and competence beliefs [Monograph]. *Journal of Vocational Behavior*, *60*, 113–172.

Tracey, T. J. G., & Rounds, J. (1997). Circular structure of vocational interests. In R. Plutchik, & H. R. Conte (Eds.), *Circumplex models of personality and emotions* (pp. 183–201). Washington, DC: American Psychological Association.



Study Skills Instruction

Maribeth Gettinger
University of Wisconsin, Madison, Wisconsin, USA

1. Introduction
2. Importance of Study Skills
3. Three Dimensions of Studying
4. Definition of Study Skills
5. Study Skills in High- Versus Low-Achieving Students
6. Rationale and Theoretical Framework for Study Skills Instruction
7. Key Components of Study Skills Instruction
8. Four Clusters of Study Skills
9. Conclusions and Future Directions
Further Reading

GLOSSARY

cognitive modeling Teacher demonstrates the use of a study strategy while "thinking aloud" to verbalize the rationale and reasoning behind its use.

cognitive study strategies Strategies that guide students to engage in appropriate thinking about the information they are studying (e.g., self-questioning).

imitation Learner applies a study strategy in a way that approximates the teacher's demonstration.

metacognitive strategies Skills that guide students in how to select, use, and monitor study tactics.

procedural study skills Behaviors or habits that allow students to maximize the benefits of available study time (e.g., time management and material organization).

repetition-based strategies Study strategies that rely on repetition or rehearsal of information to be learned.

self-regulation Students initiate the use of a study strategy on their own and make adjustments or adapt the strategy to meet their needs.

study skills Cognitive tools associated with acquiring, recording, organizing, synthesizing, remembering, and using information.

study strategy Configuration of different tactics, deliberately selected by the learner for a particular purpose or learning situation.

study tactic Specific study technique, such as underlining, note-taking, outlining, summarizing, or self-questioning.

Study skills encompass a range of coordinated cognitive processes that enhance the effectiveness and efficiency of students' learning. Study skills contribute to success in both nonacademic and academic settings. Within an information-processing framework, study skills are grouped as repetition-based, procedural, cognitive, and metacognitive skills. Key elements of study skills instruction include cognitive modeling, guided practice and feedback, and self-regulation.

1. INTRODUCTION

Study skills are fundamental for learning and academic competence. For many learners, lack of success in school is associated with deficiencies in their knowledge and application of effective study skills. Capable students at all grade levels may experience failure and frustration in school, not because they lack ability but because they lack study skills. Despite a well-documented link between academic success and

©2004 Elsevier Inc.
All rights reserved.

effective studying, students rarely receive explicit instruction in study skills. Teachers often assume that students acquire basic study skills on their own. Most learners, however, will not become proficient at studying without systematic instruction and repeated practice.

2. IMPORTANCE OF STUDY SKILLS

There are several reasons for schools to focus on developing effective study skills among students. First, study skills are necessary for learning. According to Devine's 1987 definition, study skills are the tools associated with all components of the learning process, specifically "acquiring, recording, organizing, synthesizing, remembering, and using information and ideas found in school" (p. 5). Second, effective study skills enhance students' sense of personal control and self-efficacy. Students with good study skills feel competent and confident as they learn; they are more likely to approach schoolwork with positive expectations and attitudes. Finally, study skills are basic "life skills" that apply to many activities outside of school. Study skills go beyond helping students achieve success in a classroom; they help individuals function effectively in multiple and diverse contexts (e.g., academic, athletic, social, or vocational). The development of study skills contributes to an individual's growth as an independent, lifelong learner. In his seminal 1984 paper, Rohwer noted that "studying is the principal means of self-education throughout life" (p. 1).

3. THREE DIMENSIONS OF STUDYING

A comprehensive definition of study skills derives from an understanding of what is meant by studying. Studying is defined on the basis of three key dimensions. First, studying is skillful; it requires training and practice with specific techniques that help a learner acquire, organize, retain, and use information. Second, studying is intentional. Studying requires not only the knowledge and application of skills but also volition. The act of studying is a deliberate, conscious effort by students. It depends as much on personal efficacy and motivational factors as it does on organizational and procedural skills. Finally, studying involves a self-regulatory dimension. When individuals study, they engage in self-regulatory behaviors, such as self-monitoring. Good studiers have been shown to evaluate whether understanding is occurring and whether information is being remembered.

4. DEFINITION OF STUDY SKILLS

Given these characteristics of the process of studying, what does it mean to have good study skills? Study skills is a broad term that refers to a range of competencies that enhance the effectiveness and efficiency of learning. In 1991, Ellis *et al.* made an important distinction between a study tactic and a study strategy. A study tactic is a specific technique or skill, such as underlining, note-taking, outlining, summarizing, or self-questioning. A study strategy is a configuration of different tactics, deliberately selected by the learner for a particular purpose or learning situation. Having good study skills requires both knowledge of effective tactics and strategies to execute them in a planned sequence and for a specific purpose.

5. STUDY SKILLS IN HIGH- VERSUS LOW-ACHIEVING STUDENTS

A "think-aloud" assessment method developed by Kucan and Beck in 1997, whereby students express their thoughts and verbalize their strategies while engaged in studying, has allowed researchers to explore the cognitive processes involved in effective versus ineffective studying. Rich descriptive information gained from think-aloud assessment, as well as other methods of inquiry (e.g., interviews, observation, and self-report measures), has contributed to a profile of study behaviors that differentiate successful from unsuccessful students. Characteristic study skills of high- versus low-achieving students are summarized as follows:

High-achieving students

- Do not use one study tactic; instead, they choose from a repertoire of study tactics and select study tactics in a purposeful manner.
- Take an active role in studying and justify the use of study tactics for a given learning task.
- Allocate their attention appropriately to focus on important information, jumping forward and backward to process information effectively.
- Monitor their own studying and take action to "fix up" inaccuracies in their comprehension or change study strategies.

Low-achieving students

- Have a restricted range of tactics, and use the same, often ineffective, study approach for all learning, regardless of the task content, structure, or expectations.
- Assume a passive role in learning and rely on others to tell them how to study.
- Are unaware that they must extend effort beyond simply reading content to fully comprehend and retain information.
- Do not monitor their own studying and show little evidence of looking back, or employing "fix-up" strategies to remedy comprehension problems.

6. RATIONALE AND THEORETICAL FRAMEWORK FOR STUDY SKILLS INSTRUCTION

During the past 20 years, research has provided evidence supporting the effectiveness of study skills instruction. Through explicit training in study skills, students become more efficient, thoughtful, and independent as learners, and they perform better in school. Even students who develop study skills on their own learn to study more effectively and efficiently through explicit instruction.

Although different theoretical perspectives support the benefit of equipping students with study skills, the most applicable theory is information processing. In brief, an information-processing perspective assumes that information to be learned is manipulated by the student to enhance acquisition and retention. The level of processing, or manipulation, is affected by the type of study strategy the learner uses. The more elaborate the strategy, the deeper the level of processing. An information-processing perspective provides the theoretical framework for two broad areas of research related to study skills instruction, as described next.

7. KEY COMPONENTS OF STUDY SKILLS INSTRUCTION

In recent years, there has been a trend toward developing and evaluating teaching models that emphasize the flexible use of study tactics. Evaluations of two approaches provide evidence that study strategy training improves academic performance, especially for low-achieving students. These approaches are the Strategies Intervention Model, developed by Deschler, Schumaker, and other researchers at the Kansas Institute for Research in Learning Disabilities and Center for Research on Learning, and the Good Information Processing Approach, based on the work of Pressley and colleagues.

The most effective study skills instruction incorporates a sequence of instructional phases that proceed from social modeling to a gradual increase in self-directed functioning. First, instruction begins with breaking the study strategy into basic steps, followed by explicit instruction and frequent modeling of the strategy by the teacher. Research highlights the importance of showing students, not just telling then, how to use a study skill. This type of modeling is termed cognitive modeling because the teacher demonstrates the use of a study strategy while "thinking aloud" to verbalize the rationale and reasoning behind its use. The second phase is imitation, whereby the learner applies the strategy in a way that approximates the teacher's performance. The primary role of the teacher is to provide multiple opportunities for students to practice study skills and receive feedback and guidance. Teacher guidance is gradually reduced over time as students assume increasing responsibility for using the strategy on their own. In the third phase, referred to as self-regulation, learners use the study skill independently. Opportunities are provided for students to practice using study skills with their daily assignments. Although teachers may provide minimal guidance, the expectation is that students initiate the use of a study strategy on their own and make adjustments or adapt the strategy to meet their needs. In summary, effective study skills instruction incorporates three critical components: (i) simplification of the strategy so that the basic steps are explicit, justified, and amenable to individualization; (ii) frequent modeling or demonstration of the learning strategy; and (iii) guided instruction, practice, and feedback to help students move from teacher dependence to self-dependence in judging a strategy's utility and potential effectiveness.

8. FOUR CLUSTERS OF STUDY SKILLS

In 1985, Weinstein and Mayer offered a useful framework for classifying study skills within an information-processing perspective. Specifically, study strategies are grouped on the basis of the degree of manipulation or level of processing of the information to be learned.

Using this framework, four clusters of study skills are identified: repetition-based, procedural, cognitive, and metacognitive.

The most basic study skills are repetition-based. Rehearsal strategies are most useful when storing small bits of information for the short term (e.g., studying 15 weekly spelling words) or when the content being studied is used frequently (e.g., multiplication facts). From an information-processing perspective, repetition-based study skills, although easy to learn and apply, afford minimal processing of content. Repetition strategies may be augmented to promote deeper processing of information. In 1993, Levin reviewed 20 years of research supporting the use of mnemonic devices, such as mental imagery, to enhance the effectiveness of repetition-based study skills, particularly for long-term retention and acquisition of complex information.

Procedural study skills encompass the behaviors or habits that allow students to maximize the benefits of their available study time. Several competencies underlie the development of procedural study skills, including time management, material organization, and development of schedules or consistent study routine. Organizational routines and study procedures are most effective when they are personalized by having students construct their own study plans. Although research has failed to document significant benefits for any single study routine over another, such as SQ3R (Survey, Question, Read, Respond, Review), what does contribute to positive outcomes is the consistency with which a study routine is implemented and the extent to which it is personalized or adapted by individual learners.

The goal of cognitive study strategies is to guide students to engage in appropriate thinking about the information they are studying. From an information-processing orientation, the more knowledge students have about content, the more likely they are to think about, understand, and remember it. Thus, studying is enhanced when new material is meaningful to learners and integrated with their existing knowledge. In addition, information that is stored in a network of connected facts and concepts, called schemata, is more easily learned and retained. Cognitive-based study strategies are designed to help students activate background knowledge prior to studying a topic; connect new information or concepts to what they already know; and develop new schemata, when necessary, to integrate content to be learned. Specific study skills that allow students to achieve these cognitive goals include the use of cognitive organizers or semantic maps, which are graphic representations of the interrelatedness of key ideas; question generation and summarization to activate and make connections with prior knowledge; and peer-assisted learning approaches, including tutoring, mentoring, and cooperative learning.

Whereas cognitive-based study skills relate to how learners process information, metacognitive strategies relate to how students select, monitor, and use strategies. Students with well-developed metacognitive skills are able to adjust their studying according to varying task demands; they know how to study effectively; and they understand which study strategies to deploy, monitor their studying, and allocate study time wisely. In contrast, students with weak metacognitive skills are often disorganized in their studying and have limited understanding about what to do or how to proceed with studying. Research reported by Deshler *et al.* in 1996 demonstrated that explicit training in self-questioning (e.g., "Do I understand the material I am studying?" "Should I reread or revise my study strategy?") can significantly improve students' metacognitive abilities and task performance.

9. CONCLUSIONS AND FUTURE DIRECTIONS

Several key principles are important to keep in mind when designing study skills instruction. First, students must recognize the need to implement a specific study strategy. Not all strategies are appropriate for all study tasks. For example, the most effective approach for studying spelling words is likely to be different from an effective approach for studying for a history test. Similarly, any single study tactic will require some personalization on the part of the students. In developing an awareness of different strategies, students should be encouraged to explain the appropriateness of a particular study approach for an assigned task. Routine or rote application of a study technique does not ensure effective studying. Note-taking is a good example. Students use a note-taking tactic in different ways. Some students routinely copy parts of a text verbatim and process information only superficially; others use their own words when they take notes or write notes that elaborate on the text and promote a deeper processing of information. Some students may take notes randomly; others may take notes strategically, selecting information that is important, relevant, or worthy of extra processing. How students take notes and what they do with a note-taking tactic affect how

successful their studying will be, not the activity of note-taking itself. Therefore, students need training or direction in how to use a technique effectively and to employ one that best fits the task they are studying.

In summary, acquiring effective study skills is important for academic success. Problems with studying can contribute to academic failure and inability to profit from even the most optimal instruction. To be effective studiers, student must, first, have a wide arrange of evidence-based strategies at their disposal and, second, know where, when, and how to apply these strategies. Research has demonstrated that study skills can be taught at all grade levels, thus highlighting the need for an emphasis on the development and maintenance of study skills across the curriculum for all grade levels.

See Also the Following Articles

Learning ■ Learning Styles and Approaches to Studying ■ Teaching Effectiveness ■ Transfer of Learning

Further Reading

Deshler, D. D., Ellis, E. S., & Lenz, B. K. (1996). *Teaching adolescents with learning disabilities: Strategies and methods*. Denver, CO: Love Publishing.

Devine, T. G. (1987). *Teaching study skills: A guide for teachers*. Boston: Allyn & Bacon.

Ellis, E. S., Deschler, D. D., Lenz, B. K., Schumaker, J. B., & Clark, F. L. (1991). An instructional model for teaching learning strategies. *Focus on Exceptional Children*, *23*, 1–24.

Gettinger, M., & Seibert, J. K. (2002). Contributions of study skills to academic competence. *School Psychology Review*, *31*, 350–365.

Hoover, J. J., & Patton, P. R. (1995). *Teaching students with learning problems to use study skills: A teacher's guide*. Austin, TX: Pro-Ed.

Kucan, L., & Beck, I. (1997). Thinking aloud and reading comprehension research: Inquiry, instruction, and social interaction. *Review of Educational Research*, *67*, 271–299.

Levin, J. R. (1993). Mnemonic strategies and classroom learning: A twenty-year report card. *Elementary School Journal*, *94*, 235–244.

Pressley, M., Woloshyn, V., & Associates. (1995). *Cognitive strategy instruction that really improves children's academic performance* (2nd ed.). Cambridge, MA: Brookline.

Robinson, F. P. (1970). *Effective study* (4th ed.). New York: Harper & Row.

Rohwer, W. D. (1984). An invitation to an educational psychology of studying. *Educational Psychologist*, *9*, 1–14.

Scheid, K. (1993). *Helping students become strategic learners: Guidelines for teaching*. Cambridge, MA: Brookline.

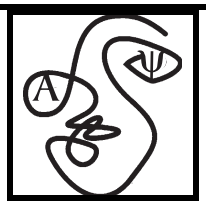
Schunk, D. H., & Zimmerman, B. (Eds.). (1998). *Self-regulated learning: From teaching to self-reflective practice*. New York: Guilford.

Strichart, S. S., Mangrum, C. T., & Iannuzzi, P. (1998). *Teaching study skills to students with learning disabilities, attention deficit disorders, or special needs*. Boston: Allyn & Bacon.

Weinstein, C. E., & Mayer, R. F. (1985). The teaching of learning strategies. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 315–329). New York: Macmillan.

Winne, P. H., & Hadwon, A. F. (1997). Studying as self-regulated learning. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 234–256). Hillsdale, NJ: Erlbaum.

Wood, E., Woloshyn, V. E., & Willoughby, T. (Eds.). (1995). *Cognitive strategy instruction for middle and high school*. Cambridge, MA: Brookline.



Subjective Culture

Harry C. Triandis
University of Illinois at Urbana–Champaign, Champaign, Illinois, USA

1. Introduction
2. Elements of Subject Culture
3. Conclusion
Further Reading

GLOSSARY

emic Refers to culture-specific associations.
etic Refers to a quality that is universal.
loose culture A quality of a group in which deviation from the group norms is tolerated.
tight culture A quality of a group in which conformity to the group norms is strictly enforced.

A broad definition of culture is that it is the human-made part of the environment. According to this definition, culture can be categorized as material and subjective culture. Material culture consists of such elements as dress, food, houses, highways, tools, and machines. Subjective culture is a society's characteristic way of perceiving its social environment. It consists of ideas about what has worked in the past and thus is worth transmitting to future generations. Language and economic, educational, political, legal, philosophical, and religious systems are important elements of culture. Ideas about aesthetics and how people should live with others are also important elements. Most important are unstated assumptions, standard operating procedures, and habits of sampling information from the environment.

1. INTRODUCTION

Subjective culture is a shared pattern of beliefs, attitudes, norms, role perceptions, and values. Thus, the first step when studying culture is to ask whether the element of culture is shared. The next step is to determine whether the shared responses correspond to a specific language, a time period, and a geographic region. Usually, people who share a language dialect at a particular time and place are members of the same culture.

Subcultures emerge because people share other elements, such as gender, physical type, neighborhood, occupation, standard of living, climate, and resources. For example, lawyers throughout world share some elements of subjective culture. Japanese lawyers have a subculture that differs from that of other lawyers as well as from general Japanese culture. A nation consists of thousands of cultures, but many of these cultures have common elements.

Since there are many elements of subjective culture, a strategy is required to study them most economically. Focusing on some of the elements of subjective culture is one way to proceed economically.

2. ELEMENTS OF SUBJECT CULTURE

2.1. Categories

Much can be learned about a culture by analyzing the categories that people use. Indigenous psychologists

©2004 Elsevier Inc.
All rights reserved.

study the meaning of specific words, such as the Japanese meaning of *amae* or the Greek meaning of *philotimos*. Such words do not have a corresponding meaning in other languages, so they tell much about a particular culture. Amae means something like expecting another person to indulge you. It is the kind of feeling one can find between a mother and a child. Pilotimos literally means friend of honor and is a common adjective used by traditional Greeks to describe themselves. It can be translated into "a person who does very frequently what family and friends expect done."

A category can be identified by noting that people give the same response to discriminably different stimuli. Our eyes, for instance, are capable of discriminating 7.5 million different color stimuli. However, we do not have millions of words that refer to color. In fact, most of us get along well with a few dozen color names.

When studying categories, we discover that some cultures have many words for a particular domain, and others have few words. For instance, the Eskimo have many words for snow; also, we have many words for cars (e.g., Ford, VW, Dodge, Toyota, truck, and vehicle). This tells us at once that the Eskimo deal with snow a lot, and we deal with cars a lot. Categories can also tell us about the way people behave. For instance, the Pawnee of Oklahoma use the same word for "mother's brother's wife," "ego's wife," and "sisters of ego's wife." We note that unacculturated Pawnees have sexual relations with all these women.

2.2. Categories Have Associations

Categories are associated with other categories. Extensive work by Osgood *et al*. has shown that throughout the world people associate categories with evaluation (good, beautiful, and moral), potency (strong, heavy, and large), and activity (fast, alive, and noisy). Since these associations occur universally, they are called etic. The term etic refers to a quality that is universal. However, there are also emic (i.e., culture-specific) associations. When we compare cultures we need to use etic constructs, but when we describe cultures we need to use emic constructs. A metaphor may help: If we compare apples and oranges we can use etic elements such as weight, size, thickness of skin, and price. However, obviously one does not learn much about such fruit with this kind of information. One needs to learn about apple flavor and orange flavor, apple texture and orange texture, and the like. These are emic qualities. Therefore, when comparing cultures etic qualities are used, but when describing cultures emic qualities are needed.

We can learn a good deal about a culture by examining its emic associations. For instance, some research has indicated that in South Korea "democracy" and "socialism" are strongly associated, but this is not the case in the United States. This tells us something about the political culture of South Korea.

2.3. Beliefs

Categories are linked to each other in other ways as well. For instance, "If a relative asks for help, you must give it" is a strong belief in some cultures, but in other cultures it is not. People in the latter cultures consider more complicated ideas, such as "Is the asking legitimate?" or "Do I like this relative?"

2.4. Attitudes

Attitudes are ideas charged with affect (emotion) predisposing action. Any category can be the core of an attitude.

2.5. Norms

Norms are ideas about behavior expected of members of a group or culture. In some cultures, called tight, people are expected to behave exactly as specified by norms. In other cultures, called loose, one can deviate from norms. In tight cultures one is punished if one does not behave according to the norms. In loose cultures a person is less likely to be punished. Punishment in some cases is very severe, whereas in other cases it is quite lenient.

When a culture is homogeneous and people are very interdependent and can be supervised closely, the culture is usually tight. When a culture is under the influence of many other cultures, or when people are not too interdependent or supervision is difficult (e.g., people live far from each other), it is more likely that the culture will be loose. Cultures are tight or loose in different domains. For instance, American culture in 2000 was very tight about writing bad checks but rather loose about who one has as a roommate. Nevertheless, across domains cultures tend toward tightness or looseness. Thai culture is loose, American culture is in-between, Japanese culture is rather tight, and theocracies such as the Taliban culture of Afghanistan are very tight.

2.6. Roles

Roles are a special category of norms. Roles are ideas about the correct behavior of people who hold a position in a social group. For example, foreman, father, aunt, or sister are roles. Roles include both prescriptive elements (e.g., fathers should advise and protect their daughters) and proscriptive elements (e.g., fathers should not hit their daughters).

2.7. Tasks

A sequence of behaviors can be defined as a task. For example, passing a law can include a series of actions. In different cultures, different sequences of actions will correspond to such a task.

2.8. Values

Values are conceptions of the desirable state of affairs. Schwartz, who has studied them extensively in many countries, defined them as beliefs that pertain to desirable states or behaviors, that transcend specific situations, that guide the selection or evaluation of behavior and events, and that are ordered by relative importance. To collect data, he uses a concept, such as "freedom," and asks people to rate its importance as "a guiding principle in my life."

3. CONCLUSION

Triandis discusses subjective culture more extensively. He also provides a discussion of methodological problems and solutions when studying subjective culture. Various chapters give examples of intensive studies of several of the elements of subjective culture.

See Also the Following Articles

Attitudes ■ Motivation and Culture

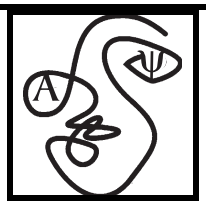
Further Reading

Kluckhohn, C. (1954). Culture and behavior. In G. Lindzey (Ed.), *Handbook of social psychology* (Vol. 2, pp. 921–976). Cambridge, MA: Addison-Wesley.

Osgood, C. E., May, W., & Miron, M. (1957). *Cross-cultural universals of affective meaning*. Urbana, IL: University of Illinois Press.

Schwartz, S. H. (1992). Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. In M. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 25). New York: Academic Press.

Triandis, H. C. (1972). *The analysis of subjective culture*. New York: Wiley.



Successful Athletic Careers

Paul Wylleman and Marc Theeboom
Vrije Universiteit, Brussels, Belgium

David Lavallee
Loughborough University, Loughborough, United Kingdom

1. Nature and Conceptualization of the Athletic Career
2. Theories and Models Related to Career Development of Athletes
3. Career Intervention Programs for Athletes
4. Conclusion
Further Reading

GLOSSARY

athletic career A succession of stages and transitions that includes an athlete's initiation into and continued participation in organized competitive sport and that is terminated with the athlete's (in)voluntary but definitive discontinuation of participation in organized competitive sport.

career program An integrated and comprehensive combination of workshops, seminars, educational modules, and individual counseling and/or a referral network providing individualized or group-oriented multidisciplinary support services to elite athletes with regard to athletic participation, developmental and lifestyle issues, and educational and vocational development.

non-normative transition An idiosyncratic transition that is generally unpredicted, unanticipated, and involuntary and that does not occur in a set plan or schedule but rather is the result of important events that take place in an athlete's life and to which she or he responds.

normative transition A transition that is part of a definite sequence of age-related biological, social, and emotional events or changes and that can be said to be generally related to the socialization process and to the organizational nature of the setting in which an athlete is involved.

successful athletic career The athletic career of an athlete who, in comparison with her or his age group peers, achieves and sustains a consistent level of athletic excellence at the national, continental, and/or world levels.

transition Generally results from one event, or a combination of events, perceived by an athlete to bring about personal and social disequilibria that are beyond the ongoing changes of everyday life and that cause a change in the athlete's assumptions about herself or himself.

A successful athletic career refers to the athletic career of an athlete who achieves and sustains in a consistent manner a level of athletic excellence in international competitions (e.g., continental or world level). Although research on the development of the athletic career during the late 1960s to early 1970s originally focused on anecdotal evidence of the process of retirement from high-level competitive and professional sports, it has become a well-delineated and growing topic of study among the sport psychology community. Research reveals that the athletic careers of elite athletes consist of succeeding normative stages, including the initiation, development, mastery, and discontinuation stages. Athletes progress in their athletic development when they cope successfully not only with the characteristics of each stage but also with the demands of the transition from one stage to another. Furthermore, athletes must cope with the demands of the stages and transitions that occur concurrently with those at the athletic level in their psychological, psychosocial, academic, and/or vocational development. Career intervention programs

©2004 Elsevier Inc.
All rights reserved.

are implemented to support athletes in optimizing and furthering their holistic development.

1. NATURE AND CONCEPTUALIZATION OF THE ATHLETIC CAREER

The academic interest in the athletic career has clearly known an increase during the past two decades. Whereas only 20 works pertaining to this area were published before 1980, more than 270 works related to the career of the athlete could be identified at the time of this writing. In addition, an inspection of the proceedings of the recent world and European sport psychology congresses reveals that nearly 10% of the program was dedicated to topics related to the athletic career. Moreover, for more than a decade, an international special interest group has focused on the exchange of information on applied and investigative work in the area. Finally, governing bodies and sport institutes around the world have developed a number of intervention programs to assist athletes in their career development.

Empirical research on the career development of elite athletes emanates from studies conducted during the 1960s and 1970s on the sport career termination among elite and professional athletes. The earliest studies revealed that athletes retiring from elite sports experienced a wide range of psychological, interpersonal, social, and financial problems initiated by a range of negative or even traumatic experiences. These studies tended to center almost exclusively on the dysfunctional issues (e.g., depression, alcoholism) experienced by retiring athletes and on male professional athletes in specific (i.e., professional) sports. However, research data gathered during the 1980s not only challenged the widespread assumption that sports retirement is inherently stressful but also led researchers to suggest that the athletic career termination should be seen as a transitional process rather than as a singular event. The increased understanding of the career termination process also motivated sport psychologists during the 1990s to broaden their attention to the occurrence of other career transitions faced by athletes. This shift toward a career development perspective was also enhanced by the integration of conceptualizations emanating from research into youth sport participation and the development of talented and expert performers. Recently, research into the development of the athletic career has taken a life span perspective by focusing not only on athletes' development from "beginning to end" at an athletic level but also on other domains of their lives (e.g., academic, psychosocial, professional).

In sum, research on the career development of elite athletes has been seen to evolve during the past four decades or so from studies on the sport career termination to studies taking a holistic, life span, multilevel approach to the sports and postsports careers of elite athletes.

2. THEORIES AND MODELS RELATED TO CAREER DEVELOPMENT OF ATHLETES

Several theories and models have been outlined in the sport psychology literature to conceptualize the athletic career. These have been employed predominantly from the mainstream psychological literature and have been developed and revised following research conducted in the area. Comparing the "end" of the athletic career to that of retirement from the workforce or even to the process of dying, sport psychologists drew from the fields of social gerontology (i.e., the study of the aging process) and of thanatology (i.e., the study of the process of dying and death). From social gerontology, "subculture" theory asserted that prolonged social interactions among individuals lead to the development of a group consciousness and that people can be less active and well adjusted during retirement even if their situations are different from overall social norms. Although athletes have fairly distinguishable (sub)cultural characteristics, and although this theory assists in revealing potential adjustment problems experienced by athletes in ending their athletic careers, the use of this theory was questioned because retiring athletes are moving out of, rather than into, the proposed subculture.

From thanatology, the "social death" approach was used to explain how, after their sports career termination, retired athletes were treated as if they were dead even though they were still biologically alive, resulting in a loss of social functioning, isolation, and/or even ostracism. This approach failed due to, among other reasons, the obvious fact that athletic retirees continued functioning in society, albeit in a different social role. A more popular use of thanatology involved describing athletic retirement in the series of stages experienced when facing death: denial and isolation

(where athletes initially refuse to acknowledge the inevitability of their career termination), anger (where retiring athletes become disturbed at the overall changing situation), bargaining (where they try to negotiate for a lengthened career in sport), depression (where they experience a distress reaction to retirement), and acceptance (where they eventually come to accept their career transition). Notwithstanding their intuitive appeal, thanatological models were criticized for the lack of an analogy between terminal illness and career termination.

More recently, models of transition have also been adopted as alternative frameworks for working with athletes in transition. These models characterize retirement from sport as a process and propose that the attributes of these athletes, cognitive appraisals of the career termination, and characteristics of the postretirement environment all interact throughout a sports career transition. Transition frameworks used in research with athletes include Sussman's "analytical model" and especially Schlossberg and colleagues' proposed "model of human adaptation to transition." In this latter model, three major sets of factors interact during a transition: the characteristics of the individual experiencing the transition, the perception of the particular transition, and the characteristics of the pre- and posttransition environments. A number of researchers also developed a sport-specific career termination model, including Taylor and Ogilvie's "model of athletic career termination." Although these models have been instrumental in stimulating research on the quality of the process of athletic retirement, their focus has remained limited to this one particular transition ending the athletic career.

Therefore, sport psychologists have adopted a developmental approach by outlining a series of predictable or "normative" transitions throughout the athletic career. This approach relates to the work by Bloom, who identified three stages of talent development (within the fields of science, art, and sport), with each stage being delineated by specific transitions. These normative transitions are part of a definite sequence of age-related biological, social, and emotional events or changes and can be said to be generally related to the socialization process as well as to the organizational nature of the setting in which the individual is involved (e.g., school, family). In the sporting domain, normative transitions include, for example, transitions from junior to senior level, from regional- to national-level competitions, from amateur to professional status, and from active participation to discontinuation from competitive sport. During these types of transition, the athlete exits one stage and enters another stage, making these transitions generally predictable and anticipated. Non-normative transitions, on the other hand, do not occur according to a set plan or schedule and are the result of important events that take place in an individual's life and to which she or he responds. For the athlete, such transitions may include a season-ending injury, the loss of a personal coach, or an unanticipated de-selection from the team. As a result, these transitions are generally unpredicted, unanticipated, and involuntary. Non-normative transitions also include those that were expected or hoped for but that did not happen (labeled "nonevents"). Not making the senior team after having made the final preselection and not being able to participate in a major championship after years of preparation are examples of nonevents.

Based on research findings showing the strong concurrent, interactive, and reciprocal nature of normative stages and transitions occurring in the athletic career and those occurring in other domains of athletes' lives, a developmental model on the athletic career that includes the normative stage transitions faced by athletes at the athletic, individual, psychosocial, and academic/vocational levels has been developed (Fig. 1). The top layer of this model reflects four stages and transitions that athletes face in their athletic development: the initiation stage (where young athletes are introduced to organized sports and where they are identified as talented athletes), the development stage (where athletes become more dedicated to their sport and where the amount of training and level of specialization is increased), the mastery (or perfection) stage (where athletes reach their highest levels of athletic proficiency), and the discontinuation stage (where elite athletes discontinue further participation in high-level competitive sport). Although this layer reflects athletes' development in elite sports, these stages can also be considered for athletes at the non elite level. The second layer of the model reflects the developmental stages and transitions that occur at the psychological level: childhood, adolescence, and (young) adulthood. The third layer of the model reflects the changes that occur in athletes' psychosocial development relative to their athletic involvement: the athletic family, peer relationships, coach–athlete relationships, marital relationships, and other interpersonal relationships significant to athletes. The final layer of the model reflects the stages and transitions at the academic and vocational levels: the transition to primary education/elementary school, the

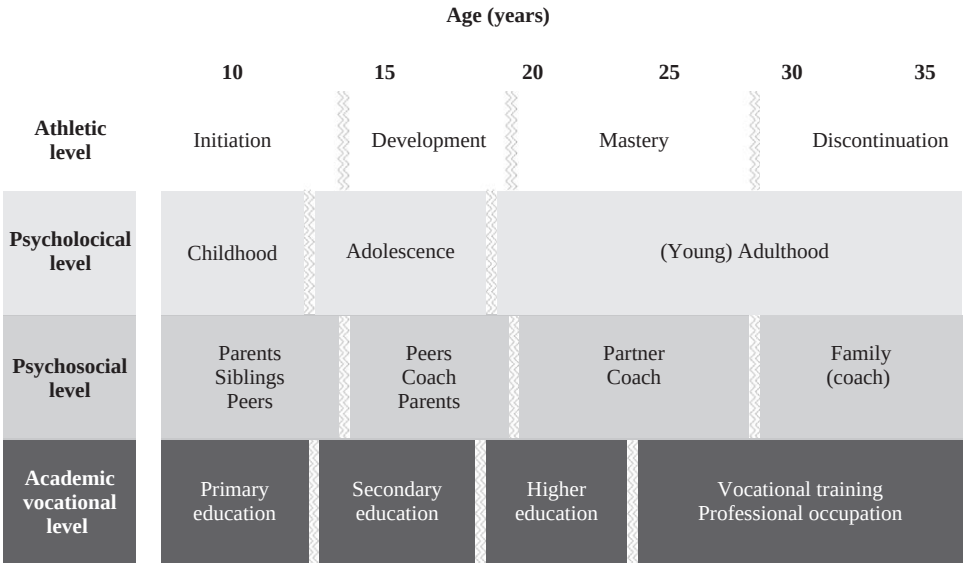


FIGURE 1 A developmental perspective on transitions faced by athletes at the athletic, psychological, psychosocial, and academic/vocational levels. Dotted lines indicate that the ages at which the transitions occur are approximations. Reprinted from Wylleman, P., & Lavallee, D. (2004). A developmental perspective on transitions faced by athletes. In M. R. Weiss (Ed.), *Developmental sport and exercise psychology: A lifespan perspective* (p. 516). Morgantown, WV: Fitness Information Technology. Reprinted with permission.

stage of secondary education/high school, the transition to higher education (college/university), and the transition to vocational training and/or a professional occupation (which may also occur at an earlier age). This model illustrates, among other things, the concurrence of transitions occurring at different levels of development (e.g., the occurrence of the transition from secondary education to higher education close to the athletic transition from the development stage to the mastery stage). This developmental model is currently used not only to determine the type of support required by talented athletes to progress in their athletic development but also to assist elite athletes in preparing for and transitioning to their postathletic careers.

In a similar vein, other researchers have identified other normative transitions. For example, Côté described the stages of sampling, specializing, investment, and mastery or performance in the development of deliberate play and practice, whereas Stambulova considered the beginning of the sport specialization, the transition to intensive training in the chosen sport, the transition to high-achievement and adult sport, the transition from amateur sport to professional sport, the transition from culmination to the end of the sport career, and the end of the sport career as predictable stages and transitions.

From an applied perspective, this developmental model enables counselors to situate and reflect on the developmental, interactive, and interdependent nature of transitions and stages faced by an individual athlete. For example, when using this model, a sport psychology consultant not only would situate a talented 16-year-old athlete in the developmental phase at the athletic level (e.g., en route to a possible selection for the national junior team) but also would consider the changes that the adolescent might be experiencing at psychological level (e.g., separation–individuation process) and psychosocial level (e.g., distancing from parents, growing influence of peers, intensifying relationship with coach) and would highlight the importance of the athlete to complete secondary education (and perhaps prepare for higher education).

However, the general focus on the identification of normative transitions should not obscure the occurrence of non-normative transitions. The unpredictable and involuntary nature of these transitions requires the development of conceptual models that include the mechanisms required by athletes to cope successfully, that is, being able to negotiate and reduce the possible negative cognitions, affects, and behaviors that may precede, accompany, or follow from these transitions.

3. CAREER INTERVENTION PROGRAMS FOR ATHLETES

During recent years, numerous career intervention programs for athletes have been developed in countries around the world (Table I). The existing programs vary in format and often include workshops, seminars, educational modules, and individual counseling. The majority of programs focus on lifestyle management and the development of transferable skills that can assist individuals in making the transition from lives in sport to postsport careers. As such, these programs provide athletes with an introduction to career planning and development by focusing on values and interest exploration, career awareness and decision making, résumé preparation, interview techniques, and job search strategies. Career development programs for athletes are primarily managed by national sports governing bodies, national Olympic committees, specific sport federations, universities, and independent organizations linked to sport settings. Although some programs address the needs of professional athletes, the majority have been developed for a much wider population (e.g., amateur players, nonelite athletes).

The Olympic Job Opportunities Program (OJOP) is an international program that has been initiated in Australia, South Africa, and the United States. The principal goal of OJOP, which is sponsored by a private company, is to develop and source career opportunities for Olympians and potential Olympians. Eligible athletes are either current or Olympic-caliber athletes who must be certified as such by their respective national federations. In addition to providing direct employer contacts and identifying job positions, OJOP provides career analysis services, personality aptitude testing, and interview skills training.

TABLE I
Selected Overview of Career Transition Programs

Program	*Institution*	*Country*
Athlete Career and Education (ACE) program	Australian Institute of Sport	Australia
Career Assistance Program for Athletes	U.S. Olympic Committee	United States
Olympic Athlete Career Center–National Sports Center	Olympic Athlete Career Center	Canada
Olympic Job Opportunities Program	Australian Olympic Committee	Australia

A grant from the U.S. Olympic Foundation was awarded to OJOP in 1988 to create a program to assist elite athletes in coping with the transition out of active sport competition. Following a survey of approximately 1800 Olympic and other elite-level athletes, the U.S. Career Assistance Program for Athletes (CAPA) was established to introduce individuals to the career development process while they were still competing. This program focused on increasing athletes' sense of personal competence through understanding and identifying transferable skills. A number of one-day workshops were organized around three main themes: managing the emotional and social impact of transitions, increasing understanding and awareness of personal qualities relevant to coping with transitions and career development, and introducing information about the world of work. Although these workshops were well received by the athletes, funding for the CAPA program was terminated in 1993.

One of the first career development programs to be developed for athletes was the Canadian Olympic Athlete Career Center (OACC). The OACC was launched in 1985 as part of the Canadian Olympic Association (COA) following a series of needs-based surveys conducted in 1983 and 1984. The original center, which was based in Toronto, had a mandate to assist athletes through the transition process to second careers, primarily through career and education planning. Athletes who had achieved approved rankings by way of their performances at Olympic, Commonwealth, and Pan-Pacific competitions had access to the OACC. The following career development services were initially offered through the OACC:

1. Clarification of career planning needs, self-assessment, aptitude/interest assessment to assist in identification of specific occupations of interest, decision making, and action planning skills were included.
2. Booklets were written and made available to all eligible athletes on the topics of curriculum preparation, interview preparation, job search techniques, and information interviewing. Retirement planning focused on what athletes should expect during the adjustment period. Transition workshops and peer support groups were established to help athletes deal effectively with the career transition process.
3. The COA provided reference letters of support and personalized business cards bearing the COA logo for networking purposes.

4. A "shadow" program was developed in 1990 to provide athletes with the opportunity to explore career options by shadowing professionals in the field of their choice.

During recent years, the COA has initiated a reorganization process that has resulted in an increase in the number of centers operating in Canada and in additions to the career development services provided. Currently, OACCs are operating in Calgary, Montreal, Ottawa, Toronto, and Vancouver, with each center employing consultants to work with athletes living in or around these locations. The standardization of service provision across centers is ongoing and is influenced by the need and utility of the services for the athlete population in each region.

The Athlete Career and Education (ACE) program was developed in 1990 in Australia by the Victorian Institute of Sport and was later amalgamated with the Lifeskills for Elite Athletes Program (SportsLEAP). SportsLEAP was originally established in 1989 by the Australian Institute of Sport based on the results of a needs-based survey following the 1988 Seoul Olympics. This program was highly successful, but the scope and content of each program varied according to athlete demand in each state institute/academy of sport. Thus, a decision was made to integrate SportsLEAP and the ACE program (under the ACE name) to form a national program in Australia.

The overall objective of the ACE program is to assist athletes in balancing the demands of their sporting careers while enhancing their opportunities to also develop their educational and vocational skills. A major component of the program is to assist individuals in developing career plans that integrate both sporting and nonsporting components. The philosophy is to create an environment where athletes can be encouraged to be independent and self-reliant and to have a capacity to meet the demands associated with elite sport. To be eligible for assistance in Australia, athletes must be scholarship holders with the Australian Institute of Sport, or with state institutes/academies of sport, or Olympic athletic program participants. ACE managers and advisers are employed in each state institute/academy of sport, and a national manager coordinates the program. The following services are provided through the ACE program:

1. Individual athlete assessments are used to provide a structured process in which to assess individual athletes' educational, vocational, financial, and personal development needs.
2. More than 30 personal development training courses are structured to assist athletes in meeting their sporting, educational, and career aspirations. Many of these courses provide individuals with nationally accredited competency-based education programs.
3. A nationally consistent career and education planning process is employed to enable athletes to manage their own individual vocational requirements. Career development is provided in the form of direct assistance in finding employment through career advice, training paths, and vocational training. Secondary and tertiary education support is provided through networking with individuals in secondary schools, universities that can offer unit or course selection advice, and assistance in negotiating appropriate academic and residential arrangements (e.g., quiet halls of residence, appropriate and sufficient dietary provisions at appropriate times, distance and online learning opportunities).
4. A transition program provides career and education guidance for elite athletes who are transitioning to postsport careers.
5. Training opportunities and supervised practice for ACE staff are provided through a graduate certificate in athlete career and education management.
6. Direct athlete needs-based assessments provide a structured process to assess athletes' eligibility for support.

ACE services are available to more than 3000 elite-level athletes throughout Australia. In addition, the U.K. Sports Institute initiated an ACE U.K. Program across England, Scotland, Wales, and Northern Ireland.

The worldwide need for services and programs that optimize athletes' development is recognized in the establishment of (inter)continental forums (e.g., International Athletes' Services Forum, the European Forum of Lifestyle Management) where relevant career and life balance issues, interventions, and programs are shared and discussed by athlete service providers and career advisers to elite athletes and teams.

4. CONCLUSION

To better appreciate how an athlete's sporting career develops, it is essential to understand not only the nature and conceptualization of the athletic career but also the theories and models related to career development. The athlete can develop a successful athletic career if she or he is able to cope with the different normative athletic and nonathletic stages

and transitions. Therefore, it is important to consider each athlete in terms of her or his athletic, psychological, psychosocial, academic, and/or vocational development and to be aware of the reciprocal influence that occurs among these various developmental contexts. As such, the existing career intervention programs for athletes around the world focus on supporting individuals through the stages and transitions that they experience both during and after their sporting careers. Applying the knowledge and conceptualization of the non-normative athletic and nonathletic transitions that affect the development of athletes' sport careers will further the development of these programs.

See Also the Following Articles

Goal Setting and Achievement Motivation in Sport ■ Intrinsic and Extrinsic Motivation in Sport ■ Psychological Skills Training in Sport ■ Self-Confidence in Athletes

Further Reading

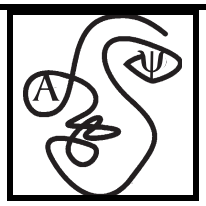
Lavallee, D., Gorely, T., Lavallee, R. M., & Wylleman, P. (2002). Career development programs for athletes. In W. Patton, & M. McMahon (Eds.), *Career development programs: Preparation for life long career decision making* (pp. 125–133). Camberwell: Australian Council for Educational Research Press.

Lavallee, D., & Wylleman, P. (Eds.). (2000). *Career transitions in sport: International perspectives*. Morgantown, WV: Fitness Information Technology.

Taylor, J., & Ogilvie, B. (2001). Career termination among athletes. In R. N. Singer, H. E. Hausenblas, & C. M. Janelle (Eds.), *Handbook of sport psychology* (pp. 672–691). New York: John Wiley.

Wylleman, P., & Lavallee, D. (2004). A developmental perspective on transitions faced by athletes. In M. Weiss (Ed.), *Developmental sport and exercise psychology: A lifespan perspective* (pp. 507–527). Morgantown, WV: Fitness Information Technology.

Wylleman, P., Lavallee, D., & Alfermann, D. (Eds.). (1999). *Career transitions in competitive sports*. Biel, Switzerland: European Federation of Sport Psychology.



Suicide Intervention in Schools

Scott Poland

Cypress–Fairbanks Independent School District, Houston, Texas, USA

1. Prevalence of Youth Suicide
2. Risk Factors for Youth Suicide
3. Most Common Misperceptions about Youth Suicide
4. Most Common Methods of Suicide
5. Role of Schools in Suicide Prevention
6. What if the Death Was a Suicide?
7. Status of State and Individual Initiatives
8. Suicide Prevention National Resources
9. Summary

Further Reading

GLOSSARY

emergency notification form A form to be signed by a suicidal student's parents that acknowledges their being notified of the suicidal emergency and that provides referral information for local mental health services.

no harm contract An agreement to be signed by a suicidal individual that promises that the individual will hot harm himself or herself and that provides alternatives for the individual in addition to suicide and crisis hotline numbers.

postvention A series of carefully planned steps to manage emotionality after a suicide, with special care being taken to be truthful but not to glorify the suicide victim.

precipitating event An emotionally troubling event that occurs just prior to a suicide attempt; this event causes the suicidal individual to act on a previously thought-out plan to attempt suicide.

suicidologists Professionals who study the problem of youth suicide and develop prevention strategies.

Youth suicide has been a constant problem for schools over the past three decades or so, with many researchers stating that it remains at or near an all-time high. Suicidologists have made a number of recommendations for school personnel, but suicide prevention has not been a primary focus in many districts and there are few or dwindling resources directed toward prevention. In 1989, Poland outlined the role of suicide prevention in the schools and identified many practical strategies and resources to assist schools in prevention efforts. In 2001, Portner found that few school districts have formal suicide prevention/intervention policies and that few counselors and psychologists have been prepared to deal with suicidal students. Portner concluded that progress in youth suicide prevention has been slow, with most prevention efforts taking place only after one or more youth suicides occurred in the same community or school. There is debate over what information should be presented to students about suicide prevention and whether the warning signs of suicide should be taught in key classes such as health. The American Association of Suicidology (AAS) has excellent resources available for schools on both prevention and postvention; however, very few schools take advantage of these resources. Surviving parents, such as those who founded Yellow Ribbon and those in Texas who have been working to convince the state legislature to pass laws so that schools have policies, procedures, and training, are making a difference; however, progress has been slow. A great

©2004 Elsevier Inc.
All rights reserved.

deal of national attention has been focused on school shootings. A U.S. Secret Service and U.S. Department of Education report on its study of 37 school shootings over several decades emphasized that the majority of school shooters were suicidal. An Associated Press article described a recent example in which a fifth-grade boy in Wellsboro, Pennsylvania, who went into a school restroom and committed suicide after his friend decided not to participate in their planned school shooting.

1. PREVALENCE OF YOUTH SUICIDE

The suicide rate for high school students ages 15 to 19 years has remained relatively stable during the past decade or so, whereas the suicide rate for young people ages 10 to 14 years has increased more than 100% during the same time period. Citing a "silent epidemic," Portner in 2001 noted that 1 in 13 teenagers had already made some form of suicide attempt. Statements that the suicide rate of young people has at least tripled since the 1950s and that it remains at or near an all-time high, are, unfortunately, very accurate. Suicide rates by age and state are available from the Centers for Disease Control and Prevention (CDC) and are typically given per 100,000 individuals in a particular age group. In the United States, suicide rates are highest in western states that have a high percentage of Native Americans, and Alaska currently has the highest rate of all states. Professionals in suicidology acknowledge that suicide rates are undoubtedly underestimated due to underreporting and the unknown causes of single-car fatal accidents. The United States has a relatively high suicide rate for young people, with suicide consistently ranking second or third as the leading cause of death among adolescents. In 2000, Poland and McCormick cited figures indicating the United States had the highest suicide rate of 26 industrialized nations that were studied. However, suicides among children under 10 years of age is rare in the United States (typically fewer than 50 suicide deaths per year), and the primary reasons given for this are as follows:

- Children are more involved with their parents.
- Children are not in the advanced intellectual level of development and, therefore, are less likely to feel alienated and lonely.
- Children are involved in less substance abuse and have less access to guns.

Native American and Caucasian male youth die by suicide most often, whereas Caucasian girls attempt suicide most often. The suicide rate for African American boys has increased the most dramatically during the past decade or so. There is also growing awareness that gay and lesbian youth are particularly at risk for suicide, although more studies are needed to clarify conflicting research. The prevalence of suicidal ideation and actions by young people necessitates that schools have policies and procedures in place to address this problem, and all school personnel must be aware of the risk factors.

2. RISK FACTORS FOR YOUTH SUICIDE

It is very important to examine the risk factors that are most involved with youth suicide. Awareness of these factors will assist school personnel in their important prevention efforts. Risk factors identified by research fall into five key areas: psychopathology, family factors, biological factors, environmental factors, and situational crises.

2.1. Psychopathology

The majority of children and adolescents who die by suicide had experienced an emotional or mental disorder, with the most common ones being affective disorder and substance abuse. Major depression and a previous suicide attempt are significant risk factors for girls, whereas a previous suicide attempt, depression, disruptive behavior, and substance abuse are significant risk factors for boys. Intoxication is involved in approximately 50% of all suicides, and alcoholics are six times more likely to die by suicide than are nonalcoholics. Researchers further emphasize the comorbidity factors involving depression, conduct problems, and alcohol abuse, especially among high school students.

2.2. Family Factors

The most significant characteristics are a family history of suicide and medical and/or psychiatric illness. In 2002, Davis and Brock found that other significant factors were economic stress, family strife, and family loss. Suicidal adolescents experience more parental divorces and report more conflict and less cohesion in their families.

2.3. Biological Factors

Evidence of serotonergic dysfunction in adult suicide attempters and completers has been found, and research continues in this area. The identification of neurochemical factors might be helpful in youth suicide prevention at some point in the future.

2.4. Environmental Factors

The strongest situational factor in the United States is the presence of a firearm. Approximately 60% of all suicides across all ages involve firearms. Many professionals believe that the most effective prevention strategy is to remove guns from ready access of troubled teens. It is noteworthy that states that have enacted strict gun laws have lower suicide rates than, for instance, western states with liberal gun laws. School personnel must not hesitate to inquire directly, during discussions with suicidal students and their parents, about gun availability and to make strong recommendations about locking up or removing guns from the homes of suicidal students.

2.5. Situational Crises

Interpersonal crises are the most common precipitating events to adolescent suicide attempts, with the breakup of romances and arguments with parents being the most common ones. Other significant precipitating events include loss, humiliation, and discipline or grade problems. School personnel and parents must be alert to precipitating events that may cause young persons to act on their previously thought out suicidal plans.

3. MOST COMMON MISPERCEPTIONS ABOUT YOUTH SUICIDE

School personnel often have many misperceptions about youth suicide, and it is important that they have accurate information. In 1999, Poland and McCormick debunked the following common myths about suicide:

Myth: Young people who talk or write about suicide do not commit suicide.

Fact: These young people are at risk and crying out for help.

TABLE I
Situational Crises

- Loss (Today's youth seem to possess more avenues to loss in their lives. Among these are loss of a loved one, losses suffered when families break up, losses caused by transience and relocation, and losses of self-esteem or romance. Traumatic grief heightens vulnerability to suicidal ideation.)
- Disciplinary crisis (e.g., fear of incarceration or trouble at school). (The rising rate of litigation against schools might indicate a need for suicidal assessment after student suspension, expulsion, or exclusion.)
- Stressful life events (e.g., poverty, academic and peer pressures)
- Family crises (e.g., family violence, parental arguments, physical or sexual abuse)
- History of running away
- Suicide completion in the community or exposure to suicidality in others

Source. Poland and Lieberman (2002). Copyright (2002 and 1995) by the national Association of School Psychologists, Bethesda, MA. Reprinted by permission of the publisher.

Myth: Suicidal youth really want to die.

Fact: Suicidal youth are ambivalent about dying but want the pain to end and wish that someone would change the situation in a positive way so that their lives can continue.

Myth: Discussing suicide will give a suicidal young person the idea to commit suicide.

Fact: All young people are aware that suicide is an option, and discussing it will provide assurances that others care about them. It will empower them that they can save their own lives or the lives of friends by getting adult help.

Myth: Suicide is inherited or destined.

Fact: Although parental depression is a predictor of adolescent suicide, the most important variables are situational crises, as outlined in Table I.

Myth: There are usually no warning signs of suicide.

Fact: Nearly all adolescents who die by suicide gave some warning to friends or family members. Many professionals believe that these adolescents were seeking help and options besides suicide.

4. MOST COMMON METHODS OF SUICIDE

Prior to 1980, the use of pills and overdoses was the number one method of suicide among girls. The use of

guns, followed by hanging and explosions, has long been the number one method among boys. More recent female suicide methods, in order of frequency, are gunshots, pill overdoses, and hanging.

Many professionals believe that removing access to guns from a suicidal individual is essential to reducing suicides given that approximately 60% of all suicides across all ages involve a gun. Professionals recommend that school personnel inquire very directly and specifically about the presence of any guns in the homes of suicidal students. Parents who are notified by school personnel that their children are suicidal must be convinced to securely lock up their guns or remove them from their homes.

It is very important that school personnel be very aware of the most common methods of suicide so that appropriate actions can be taken to remove any available lethal means of suicide. Parents and school personnel must be alert to the precipitating events that often precede youth suicide so that appropriate interventions can take place as soon as possible to prevent the suicide from occurring. The most common precipitating events are the following:

- Severe arguments with parents
- Breakup of a romance
- Loss of a loved one
- Discipline or grade problems

5. ROLE OF SCHOOLS IN SUICIDE PREVENTION

There has been a growing awareness of the need for schools to be more knowledgeable about suicide prevention and to regularly in-service school personnel on the warning signs of suicide as well as on the importance of all school personnel working as a team to detect suicidal students and secure assistance for them. Many state lawmakers have passed legislation addressing prevention in the schools. Unfortunately, most of the legislation that has already been passed has involved recommendations rather than mandatory actions. The nationwide emphasis on high-stakes testing and academic accountability has also minimized the focus on suicide prevention as well as on other important mental health issues. For example, a school superintendent in Austin, Texas, commented that, given a choice of providing teachers with an in-service on preventive mental health topics or with one that focuses on improving test scores, he would opt for the one on test scores.

It is also important to recognize that there have been attempts to hold schools accountable for the suicides of students, and school districts have been subjected to lawsuits over the suicides of students. The question is not whether the school somehow caused the suicide (regardless of whether the suicide occurred on or off school grounds) but rather whether the school took reasonable action to prevent it from occurring. The issues, as outlined by Lieberman and Davis in 2002, are as follows:

- *Foreseeability:* Did the school have reason to foresee that the young person was at risk for suicide? Were warning signs obvious?
- *Negligence:* It is negligent to fail to secure counseling assistance for a suicidal student. It is negligent to fail to notify parents when a child is suicidal.

In 1999, Poland and McCormick reviewed a number of cases where school personnel did not follow commonsense practices of notifying parents, increasing supervision, and providing referral and follow-up services. School personnel and school districts were found by courts to be liable for monetary damages. One such case in 1997 was *Wyke v. Polk County School Board* following the death by suicide of a student who school administrators knew was suicidal. The student's parents were awarded a monetary judgment. The 1991 case of *Eidsel v. Board of Education of Montgomery County* illustrated the importance of parents being notified when school personnel suspect that a student is suicidal, even though the student himself or herself denies suicidal intent.

Administrators are in a particularly difficult situation when a student is facing suspension and expulsion consequences and states, "If you expel me, then I am going home to kill myself." There have been situations where discipline incidents precipitated suicides. Administrators must be alert to suicidal ideation for students who are involved in discipline situations and must follow procedures that include parent notification. The role of the schools in suicide intervention is outlined in Table II, and various suicide warning signs are provided in Table III. Figures 1 and 2 illustrate a sample no suicide or no harm contract and a sample parent form for notification of child's suicidal ideation, respectively.

TABLE II
The School's Role in Suicide Prevention

- *Awareness and detection:* All school personnel who supervise students (i.e., who have any form of contact with students), including administrators, teachers, counselors, psychologists, speech therapists, classroom aides, and bus drivers, should be in-serviced annually on the warning signs and what to do when they suspect a student might be suicidal. In addition, it is recommended that awareness training include personnel such as office secretaries, custodians, and cafeteria workers.
- *Assessment:* Support personnel such as counselors and psychologists often experience anxiety when working with suicidal students and need specialized training to increase their confidence level to assess the severity level of the student's suicidal behavior. It is important that support personnel placed in this role collaborate with other professionals and maintain frequent communication with the school administrators. In 1995, Poland outlined Texas legislation stating that parental permission is not necessary prior to seeing a minor when the minor is believed to be suicidal. Poland further stressed the importance of a direct inquiry with questions such as the following:

  Have you thought about harming yourself or trying to end your life?

  Have you previously attempted suicide?

  Have you thought about suicide lately?

  How would you end your life?

  The answers to these questions are essential to know whether parents must be contacted immediately and whether the student must be kept under close supervision until parents can be notified.
- A "*no harm contract*" is another essential step in suicide prevention. Copies of the signed contract should be given to the suicidal student, his or her parents, and the school administrator.
- *Parent notification:* Parents must be made aware that it is suspected that their child is suicidal. The question is not whether to call the parents but rather what to say. The conversation might need to include only statements such as that the child is depressed and has lost the joy for life or that the child has an imminent plan to commit suicide and that parents need to come to the school immediately. Documentation of parent notification is essential. The role of the school should be outlined, and referral for additional community services should be made.
- *Referral and follow-up services:* Students known to be suicidal should be referred to community agencies and private practitioners for more in-depth counseling than school employees can provide. School personnel, such as administrators, counselors, and psychologists, should monitor suicidal students closely, including frequent contact with the students, their parents, and any outside professionals providing treatment. Protective service agencies should be notified when parents do not seek counseling assistance for suicidal students.

6. WHAT IF THE DEATH WAS A SUICIDE?

In the case of a student's death, it is very important for school personnel to talk with the family to verify the cause of death and to ask permission to inform the staff and students at the deceased child's school of the circumstances. The AAS's Guidelines for Postvention are an excellent source outlining the role of the schools as well as media guidelines. These guidelines, summarized in Table IV, emphasize that being truthful about the circumstances is important and that postvention is an ideal opportunity in which to raise awareness to prevent the future suicides. Fearing that attention to a recent suicide will be perceived as glamorizing the deceased and the circumstances, schools often prefer to do very little; however, it is important to deal with survivors' loss, confusion, shock, and grief by empowering them with a resolve to prevent further suicides. A particularly troubling issue for school administrators is the question of whether to use a suicide victim's name in a memorandum to teachers and in a note to parents. The AAS guidelines specifically give examples of stating the name of the suicide victim; however, schools commonly do not provide that information, even though the cause of death is a matter of public record.

Few events in a school are more troubling than the suicide of a student. The school crisis team would greatly benefit from reviewing the AAS's guidelines for postvention that also outline the need to be very cautious about memorializing the suicide victim. School superintendents and the principal must carefully consider the requests of family and friends to memorialize the suicide victim. Permanent markers and memorial activities at school are not recommended. It is hoped that the victim's family can be convinced to donate money to suicide prevention crisis centers or to scholarship funds.

TABLE III
Warning Signs

- *Threat:* Threats may be direct (e.g., "I want to die," "I am going to kill myself") or indirect (e.g., "The world would be better without me," "Nobody will miss me anyway"). During adolescence, indirect clues could be offered through joking or through references in school assignments (e.g., creative writing, art pieces). In concrete and preoperational children, indirect clues may come in the form of acting out or violent behavior that is often accompanied by suicidal/homicidal threats.
- *Plan/Method/Access:* Suicidal thoughts must be distinguished from actual planning. The greater the planning, the greater the potential for suicide. In evaluating the suicidal potential of a student, the lethal potential and availability of the means and the level of sophistication of the plan (including the developmental level of the interviewee) must be taken into account. Familiar themes for children include running into traffic, jumping from high places, and cutting/scratching/marking their bodies.
- *Previous attempts:* Adolescent attempters are at increased risk for a variety of negative outcomes, including repeat attempts, psychiatric symptoms, and academic, social, and behavioral problems. Approximately 15% of individuals with a history of one or more suicide attempts will go on to kill themselves.
- *Final arrangements:* This behavior may take many forms. In adolescents, it might be giving away prized possessions (e.g., jewelry, clothing, journal). It seems likely that preoperational elementary children lack the cognitive skills necessary to plan for making final arrangements.
- *Depression:* When symptoms of depression include pervasive thoughts of helplessness and hopelessness, children or adolescents are conceivably at greater risk for suicide.
- *Sudden changes:* Changes in behavior, friends, or personality are warning signs.

Source. Poland and Lieberman (2002). Copyright (2002 and 1995) by the national Association of School Psychologists, Bethesda, MA. Reprinted by permission of the publisher.

In 2002, Poland and Lieberman cited literature predicting that with the occurrence of one youth suicide in a community, the chances of a second suicide increase dramatically. Teenagers in particular are more susceptible to imitating classmates who have died by suicide. The question of suicide contagion or clusters continues to be studied by the CDC.

7. STATUS OF STATE AND INDIVIDUAL INITIATIVES

The majority of states have some form of legislation addressing the problem of youth suicide. California has long been a leader in youth suicide prevention, but prevention initiatives have lost momentum in many states and are largely only recommendations. Many well-meaning school administrators have difficulty in finding the time for teacher training as well as in developing and implementing curriculum units to teach students about suicide prevention. Many of the youth suicide prevention initiatives currently under way at the state level are being driven by surviving parents.

I, _____ , agree not to harm myself.

If I am having thoughts of harming myself or committing suicide, then I will do the following until I receive help:
Get assistance from an adult.
Call the Crisis Hotline at _____.
Call the school psychologist/counselor at _____.

I understand the contract that I am signing and agree to abide by it.

_____(student signature)

_____ (school psychologist)

FIGURE 1 Sample of no suicide or no harm contract. Reprinted from Poland (1995). Copyright (2002 and 1995) by the national Association of School Psychologists, Bethesda, MA. Reprinted by permission of the publisher.

I/We, _____, the parents of _____, were involved in a conference with school personnel on _____. We have been notified that our child is suicidal. We have been further advised that we should seek some psychological/psychiatric consultation immediately from the community. School personnel have clarified the district's role and will provide follow-up assistance to our child to support the treatment services from the community.

_____ _____
(parent or legal guardian) (school personnel/title)

FIGURE 2 Sample of parent form for notification of child's suicidal ideation. Reprinted from Poland (1995). Copyright (2002 and 1995) by the national Association of School Psychologists, Bethesda, MA. Reprinted by permission of the publisher.

TABLE IV
General Suicide Postvention Guidelines for Schools

- Plan in advance of any crisis.
- Select and train a crisis team.
- Verify report of suicide from collaboration with the medical examiner, police, and family of the deceased.
- Do not dismiss school, and do not encourage funeral attendance, during school hours.
- Do not dedicate a memorial (e.g., yearbooks, tree, bench).
- Do contribute to a suicide prevention effort on behalf of the school or community.
- Do contact the family, apprise them of the school's intervention efforts, and assist with funeral arrangements.
- Do not release information in a large assembly or over intercom systems. Instead, disseminate information to faculty, students, and parents in small groups or in a classroom setting. Always be truthful.
- Follow the victim's classes throughout the day with discussion and counseling.
- Arrange for counseling rooms in the school building, and provide individual and group counseling.
- Collaborate with media, law enforcement, and community agencies.
- Emphasize the following points with media and parents:
  - Prevention is crucial.
  - No one thing or person is to blame.
  - Help is available.
  - Provide counseling or discussion opportunities for the faculty.
  - Discourage media from including front-page coverage of the suicide or including a picture of the deceased or a description of the method.
  - Discourage media from using simplistic or romantic explanation of the suicide.
  - Encourage media to include a list of suicide warning signs and specific examples of what to do to secure professional help for a suicidal individual.

Source. American Association of Suicidology (1998) and Poland and Lieberman (2002). Copyright (2002 and 1995) by the national Association of School Psychologists, Bethesda, MA. Reprinted by permission of the publisher.

School-based suicide prevention programs fall into three categories as outlined by Kalafat and Lazarus in 2002: universal interventions, selective interventions, and indicated interventions. Universal interventions are directed to an entire population and involve increasing likelihood that administrators, faculty, and peers who come into contact with suicidal youth can more readily identify them and secure help for them, an approach that is often referred to as "gatekeeper training." Teaching school personnel the warning signs and developing guidelines for schools to follow when students are identified as suicidal has been the most widely used approach to suicide prevention. An additional universal intervention is having students participate in a national suicide screening day where large numbers of students answer questionnaires that address depression and suicide. This program targets secondary student and is sponsored by the National Association of School Psychologists. The program includes as a first step an educational component where students are taught the warning signs and the importance of getting adult help for anyone believed to be suicidal. Students then fill out a questionnaire that asks questions about depression and suicide. Students who are identified as depressed or suicidal, based on their responses, are evaluated further and are provided with interventions. Typically, the screening is done 1 day each year at participating schools; thus, a limitation is that a student might become suicidal at a later point in the year without having reported suicidal behavior/thoughts on the screening date. Approximately 1500 schools have implemented this program and have reported increases in adolescent help-seeking behaviors.

Selective interventions target subpopulations that are characterized by shared exposure to risk factors such as making the transition to a new school and attending a school where a number of students have died by suicide.

Indicated interventions target specific individuals who have been identified as suicidal by screening procedures, self-referral, and/or referral by a concerned peer or faculty member.

Kalafat and Lazarus pointed out that it is important to view these as complementary approaches rather than competing approaches. They summarized the research on suicide prevention programs and emphasized that although early research results during the 1980s about the effectiveness of suicide prevention programs were mixed, more recent controlled analysis of suicide prevention programs in the schools has resulted in reduced suicidal behavior and in lives being saved. The authors stressed the need for additional research and linkage between school-based programs and community resources.

A very promising development at the national level was initiated in the U.S. surgeon general's "Call to Action to Prevent Suicide" in 1999 and in the U.S. Public Health Service's "National Strategy for Suicide

TABLE V
National Suicide Prevention Efforts

Surgeon general's "Call to Action to Prevent Suicide" (1999)
- *Awareness:* Broaden the public's awareness of suicide and its risk factors.
- *Intervention:* Enhance services and programs, both population based and clinical care.
- *Methodology:* Advance the science of suicide.

National Strategy for Suicide Prevention's Goals and Objectives for Action (2001):
- *Goal 1:* Promote awareness that suicide is a public health problem that is preventable.
- *Goal 2:* Develop broad-based support for suicide prevention.
- *Goal 3:* Develop and implement strategies to reduce the stigma association.
- *Goal 4:* Develop and implement suicide prevention programs.
- *Goal 5:* Promote efforts to reduce access to lethal means and methods of self-harm.
- *Goal 6:* Implement training for recognition of at-risk behavior and delivery of effective treatment.
- *Goal 7:* Develop and promote effective clinical and professional practices.
- *Goal 8:* Improve access to and community linkages with mental health and substance abuse services.
- *Goal 9:* Improve reporting and portrayals of suicidal behavior in the entertainment and news media.
- *Goal 10:* Promote and support research on suicide and suicide prevention.
- *Goal 11:* Improve and expand surveillance systems.

Note. Goals outlined in National Strategy for Suicide Prevention are very important. However, little funding has been attached to them, and there has been little clarification of where schools fit into these national goals.

Prevention" in 2001. These actions outlined the goals listed in Table V.

8. SUICIDE PREVENTION NATIONAL RESOURCES

School personnel are encouraged to familiarize themselves with the initiatives in their own states as well as at the national level. They should also contact local crisis hotline centers and use them as a resource in working on youth suicide prevention in the school and community. Table VI lists national resources.

TABLE VI
Suicide Prevention Resources

- American Association of Suicidology. www.suicidology.org. Many resources are available from AAS. Its members are a combination of professionals working in suicide prevention and laypersons who are suicide survivors. An annual convention during the spring.
- Yellow Ribbon Suicide Prevention Foundation. www.yellowribbon.org. Founded by parents who lost a child to suicide, this organization offers support for survivors and prevention information for schools. Local chapters are located in nearly every state as well as in several countries. The annual international convention is held in Denver, Colorado.
- Americans for Suicide Prevention. www.afsp.org.
- Suicide Prevention Awareness Network. www.spanusa.org.
- International Association for Suicide Prévention. www.med.uio.no/iasp.
- National Suicide Prévention Hotline. 1-800-SUICIDE.

9. SUMMARY

The World Health Organization's report on suicide in 2000 emphasized the following protective factors to prevent youth suicide:

- Good family relations and support
- Good relationships with teachers and peers
- Social integration and participation in school and community
- Social skills and self-confidence
- Willingness to seek help when difficulties arise

These factors are important for the success and satisfaction of all children. However, they involve a partnership/collaboration among schools, parents, and community leaders. School psychologists and other mental health professionals, such as school counselors and social workers, are in a unique position to encourage and help schools to develop and implement policies and procedures to prevent youth suicide through outlining each school employee's role in the prevention effort. It is also essential to educate the students themselves about how they can help to prevent suicides. It is hoped that the National Strategy for Suicide Prevention will result in many highly funded initiatives that will clarify the role of schools, provide more mental health services to suicidal students, and encourage schools, agencies, and organizations to work together. It is also

hoped that no more schools and communities will suffer the shock, devastation, and confusion that result from the suicides of students or staff members.

See Also the Following Articles

Behavioral Observation in Schools ■ Bullying and Abuse on School Campuses ■ Stress

Further Reading

American Association of Suicidology. (1998). *Suicide postvention guidelines: Suggestions for dealing with the aftermath of suicide in the schools*. Washington, DC: Author.

Associated Press. (2003, June 12). School shooting in Pennsylvania ends with student suicide. *Houston Chronicle, P*. A6.

Brent, D. A., & Perper, J. A. (1995). Research in adolescent suicide: Implications for training, service delivery, and public policy. *Suicide and Life-Threatening Behavior*, *25*, 222–230.

Davis, J. M., & Brock, S. E. (2002). Suicide and the schools: Incidence, theory, research, and response. In J. Sandoval (Ed.), *Crisis counseling, intervention, and prevention in the schools* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum.

Eisel v. Board of Education of Montgomery County, 597 A.2nd 447 Maryland (1991).

Kalafat, J., & Lazarus, P. (2002). Suicide prevention in the schools. In S. E. Brock, P. H. Lazarus, & S. R. Jimerson (Eds.), *Best practices in school crisis prevention and intervention*. Bethesda, MD: National Association of School Psychologists.

Lieberman, R., & Davis, J. (2002). Suicide intervention. In S. E. Brock, P. J. Lazarus, & S. R. Jimerson (Eds.), *Best practices in school crisis prevention and intervention*. Bethesda, MD: National Association of School Psychologists.

Moscicki, E. K. (1995). Epidemiology of suicidal behavior. *Suicide and Life-Threatening Behavior*, *25*, 22–35.

Poland, S. (1989). *Suicide intervention in the schools*. New York: Guilford.

Poland, S. (1995). Best practices in suicide intervention. In A. Thomas, & J. Grimes (Eds.), *Best practices in school psychology* (Vol. 3, pp. 155–166). Bethesda, MD: National Association of School Psychologists.

Poland, S., & Lieberman, R. (2002). Best practices in suicide intervention. In A. Thomas, & J. Grimes (Eds.), *Best practices in school psychology* (Vol. 4, pp. 1151–1166). Bethesda, MD: National Association of School Psychologists.

Poland, S., & McCormick, J. S. (1999). *Coping with crisis: Lessons learned*. Longmont, CO: Sopris West.

Poland, S., & McCormick, J. (2000). *Coping with crisis: A quick reference guide*. Longmont, CO: Sopris West.

Portner, J. (2001). *One in thirteen: The silent epidemic of teen suicide*. Beltsville, MD: Robins Lane Press.

Ramsay, R. F., Tanney, B. L., Tierney, R. J., & Lang, W. A. (1996). *Suicide intervention workshop* (6th ed.). Calgary, Alberta: Living Works Education.

Shaffer, D., & Piacentini, J. (1994). Suicide and suicide attempts. In M. Rutter, L. Hersov, & E. Taylor (Eds.), *Child and adolescent psychiatry* (3rd ed., pp. 407–424). London: Blackwell Scientific.

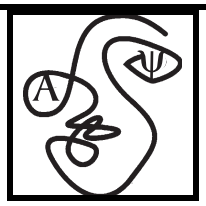
U.S. Public Health Service. (1999).*The surgeon general's call to action to prevent suicide*. Washington, DC: Author.

U.S. Public Health Service. (2001).*National strategy for suicide prevention: Goals and objectives for action*. Washington, DC: Author. Available: www.mentalhealth.org.

Vossekuil, B., Reddy, M., Fein, R., Borum, R., & Modzeleski, W. (2000).*U.S. Safe School Initiative: An interim report on the prevention of targeted violence in schools*. Washington, DC: U.S. Secret Service, National Threat Assessment Center.

World Health Organization. (2000). *Preventing suicide: A resource for teachers and other school staff*. Geneva, Switzerland: Author.

Wyke v. Polk County School Board, 11th Cir. (1997).



System Safety

Bernhard Wilpert
Berlin University of Technology, Berlin, Germany

1. Large Complex Sociotechnical Systems
2. Safety Management Strategies
3. The Role of Psychology in System Safety
Further Reading

GLOSSARY

event analysis The systematic approach to identifying factors that contributed to an event (incident, accident, or near accident) in organizations.

large complex sociotechnical systems Very large industrial installations of high complexity with a concomitant increased hazard level for the environment and people.

safety management systems The portfolio of concrete structural and organizational measures initiated by management in order to maintain safety and reliability of a given sociotechnical system.

system safety The quality of a system that allows the system to function without major breakdowns under predetermined conditions with an acceptable minimum of accidental loss and unintended harm to the organization and its environment.

Applied psychology traditionally studied safety as a problem of individuals in their respective workplaces and was concerned with the avoidance of workplace accidents. The growth of large and highly complex sociotechnical installations brought with it a drastic increase in concomitant hazards. Thus, applied psychology enlarged its scope of reference to include the safety and reliability of highly hazardous systems and began to focus on the avoidance of total system breakdowns. Applied psychology has an important role in the development and implementation of system safety strategies.

1. LARGE COMPLEX SOCIOTECHNICAL SYSTEMS

During the second half of the 20th century, the number and complexity of very large technical installations increased greatly. Along with their growth was a concomitant increase in hazardousness. Examples of these installations include those of the chemical and petrol industry, civil and military aviation, railway systems and space industry, and nuclear energy production—that is, installations in unforgiving environments in which the accumulation of high energy or toxic substances must be controlled to avoid catastrophic consequences for people and environment. Their characteristic features include the involvement of multiple actors, high levels of complexity, tightly coupled subsystems and production processes with their concomitant opaqueness and control problems, and comparatively high hazard levels. The thorough analysis of any major industrial accident (e.g., Three Mile Island, Chernobyl, Challenger, the collision of two jumbo planes in Tenerife, and the train accident in Eschede, Germany) invariably shows that it is never a single component failure or a single human error that causes the accident. Rather, it is the

©2004 Elsevier Inc.
All rights reserved.

often little understood interaction of factors from various system levels that "causes" accidents: failures on the technical (component) level, erroneous individual (operator) action, inadequate work team activities, and dysfunctionalities in the organizational and extraorganizational environment.

Münsterberg, in his seminal publication *Psychology and Industrial Efficiency*, showed in 1913 that safety in work systems has from its very beginning been a preoccupation of applied psychology. However, its focus tended to be occupational safety of individuals in their specific workplace settings. "Human factors" in the tradition of ergonomics and occupational safety were considered to refer to all characteristics at the immediate man–machine interface that enable proper and safe operations. Although recently there has been a shift in theory and practice of safety, health, and environment to more holistic approaches encompassing the whole organizational system, large-scale, highly hazardous systems pose novel challenges for psychology in terms of theory, methodological analysis, and safety-oriented intervention practices. Safety and reliability of complex systems depend on failure avoidance and effective process and production control. From the analysis of major accidents, we know that human action always plays a major role in triggering accidents as well as mitigating their consequences.

Work and organizational psychology has for a long time neglected to focus on the role of humans in safeguarding highly complex sociotechnical systems. In 1974, Graumann noted that one reason may have been the preoccupation with individual organismic perspectives at the expense of studying the interaction of humans with "the world of things." The focus of organizational psychology was occupational hazards and individual workplace accidents as influenced by individual constructs such as motivation, perception, attitude, and decision making. The world of things (i.e., complex technologies and large-scale installations) was seen to fall into the domain of engineers, who traditionally tended to focus on component failure. However, neither the exclusive individual or workplace orientation nor the limited preoccupation with technical aspects help to derive a proper understanding of the genesis of system breakdowns.

The comparative neglect of the multitude of human and physical/environmental factors in producing intended or unintended outcomes is surprising because there are at least two important theoretical strands in applied psychology that shed crucial light on these relations: Kurt Lewin's field theory and the sociotechnical systems approach of the Tavistock Institute of Human Relations. They have been developed for more than half a century, and both of these perspectives call for a comprehensive consideration of all factors that impact a given outcome under study, including the behavior of people (e.g., safety relevant actions, commitment, and questioning procedures) or systems (e.g., efficiency of goal getting, safety, reliability, and profitability). These factors may be located on various systems levels—the individual, the work group, the organization, the interorganizational, or the technological. For traditional psychology, however, such a holistic systemic thinking is rare. The focus is traditionally restricted to the individual subsystem.

Regarding system safety, using Amalberti's classification we can distinguish three safety levels in different systems:

Level 1 refers to nonprofessional domains in which the safety level is lower than 10E-3; for example, in situations such as mountaineering in the Himalayas, with a survival chance of only 30% after three expeditions.

Level 2 corresponds to systems that over time reach a safety level of 10E-5, such as street traffic, anesthesiology, and helicopter flights.

Level 3 covers systems with a safety level near or higher than 10E-6—so-called ultracomplex and ultrasafe systems such as nuclear power plants, civil aviation, and space travel.

The nature of ultrasafe, large-scale technical systems internationally brought about a paradigmatic shift in safety thinking. Traditional workplace safety concerns tried to avoid individual accidents in work settings by creating safe working conditions. Ergonomics, work medicine, work and organizational psychology were the disciplines in demand. The emergence of large-scale technical systems required that specialists in psychology, engineering sciences, and political and legal sciences cooperate in order to prevent "systems disasters"—large industrial catastrophes due to total systems failures (e.g., the Chernobyl accident). In agreement with the formulation of Roland and Moriarty, we can define system safety as a quality of a system that allows the system to function without major breakdowns under predetermined conditions with an acceptable minimum of accidental loss and unintended harm to the organization and its environment. The critical elements of this definition refer to the systemic quality of safety in the sense that all subsystems of a given organization are considered

contributors to the intended outcome: (i) safety as avoidance of system breakdown; (ii) the system's functioning is linked to predetermined performance criteria; and (iii) the system performance is measured against a socially acceptable limit of uncontrollable, if negative, side effects. One consequence of such a holistic view of system safety is that "human factors" not only relate to the immediate man–machine interface but also human action on all system levels in as much as they contribute to the critical outcome variable: safety.

2. SAFETY MANAGEMENT STRATEGIES

In 1991, Rasmussen, in a somewhat abstract manner, distinguished three different approaches to the control of safety. First, the feedback strategy reactively attempts to control relatively low-risk and unassociated hazards by empirically monitoring acceptable safety goals as measured by accident and injury rates. Such a strategy may be appropriate in traffic and construction systems, which have a slow pace of change. It is principally oriented toward learning from past events and data suggesting specific behavior adjustments.

Second, the feedforward strategy concentrates on control through enlightened design and planning of operations in the sense of analyzing probable system hazards in advance and implementing requisite safeguards into the system. The approach seems appropriate for highly hazardous systems with slow change such as railway systems, in which methods of probabilistic analyses (e.g., Probabilistic Safety Analysis) are fruitfully employed.

Finally, the feedback and feedforward strategy combined tries to learn from past events. It integrates such learning into probabilistic methods and is oriented toward evolutionary choice, which makes it particularly appropriate for ultrasafe systems such as nuclear power plants.

In a different approach, Reason identified three safety management models: a person model, which focuses on individual work-related competencies and behavior; an engineering model, with emphasis on improving technology and reducing man–machine mismatches to enhance safety; and an organizational model, in which avoidance of latent weaknesses (insufficient staff training, inadequate managerial decisions regarding division of work, and inappropriate incentive systems) in the organization are to be minimized. Applied psychology has significant contributions to make for all three of Reason's models: in the person model, action regulation theory; in the engineering model, physiological psychology, with its insights into the characteristics (possibilities and limits) of the sensory nervous system of operators that need to be matched with workplace demands in order to maximize safety; and in the organization model, leadership theories, with their implications for requisite training measures.

Different managerial safety strategies have operationally been made more precise under the term "safety management systems." The term covers a large number of concrete activities and measures, such as the outlining of a precisely defined safety policy by management, the clear distribution of responsibilities among staff and departments, the adequate allocation of resources to organizational units, application of procedures to assess operational risks, learning from operational experience through incident reporting and incident analyses, competence-enhancing training measures, monitoring of safety-related data, and internal and external audits and peer reviews.

3. THE ROLE OF PSYCHOLOGY IN SYSTEM SAFETY

System safety is not the exclusive interest of psychologists, as demonstrated by the founding of the System Safety Society in 1964, which is international in scope and has members from various professional backgrounds. However, there are five important reasons why the field of psychology should be involved in issues of system safety.

First, an important, if not the only, precondition of individual growth is the personal, bodily integrity of working people and their environment. This is the ethical mandate for psychologists to contribute to this precondition.

Second, since people are almost always involved in triggering or mitigating accidents, it seems to be self-evident that the scientific discipline that deals with human perception, competence, motivation, and action is challenged here as well. This is the mandate that follows from the self-image of the psychological discipline.

Third, no other scientific discipline is as prepared and geared to deal with many of the problems raised by highly hazardous organizations as is psychology. Psychology deals with the physiological bases of human action (e.g., information-processing capacities),

cognitive dynamics (perceptual distortions), organizational conditions of work, strain and stress, group dynamics, and risk communication.

The fourth reason concerns the enormous hazard potential implied in highly hazardous organizations. Hence, from an ecological standpoint, one must not leave the area to engineers alone.

Finally, technological development provides an important reason: The increasing complexity of new technological systems calls for an understanding of technology that transcends material objects (machines, software, etc.) traditionally understood as technology. Today, even engineers understand technology in broader terms: Technology comprises the material artifacts such as machines but also the human action facilitated by them as well as the set of human actions in which these artifacts are used. In other words, the conditions of emerging material technological systems as well as their use must be considered as technology. This is the interdisciplinary contribution of psychology that is called for.

Three examples illustrate the potential contributions of applied psychology to system safety beyond the general reasons given previously.

3.1. Event Analysis and Event Reporting

The feedback strategy of safety control suggests that learning needs to take place in highly hazardous organizations through a systematic analysis of the organization's operational experience. Individuals learn from mistakes, whereas organizations learn from incidents. Hence, it has become a generally accepted tenet in highly hazardous industries that it is necessary to thoroughly analyze, document, and report events (incidents and accidents). Given that the analysis of events is a socially accepted reconstruction of the processes leading up to the event, the analysis implies the need to conjecture about causal relations that usually goes beyond the immediately observable facts. Such attributional processes are fraught with potential biases that are well analyzed and understood by various subdisciplines of psychology, such as cognitive psychology and social psychology. Biases that result from general human attributional tendencies in the search for causal relations include the following:

- Premature hypothesis formation
- Overweighting of immediately present factors and neglect of distant ones ("out of sight—out of mind")
- Monocausal thinking
- The fundamental attribution error (causes foremost attributed to acting persons rather than situational factors)

The Research Center Systems Safety of the Institute of Psychology and Ergonomics, Berlin University of Technology developed a methodology, Safety through Organizational Learning (SOL), that is available as a computer-assisted tool. SOL is used by German and Swiss nuclear power plants as the method of choice for analyses of their events.

3.2. Training and Personnel Development

Maintaining and improving the competence of managers and staff in highly hazardous industries is a constant and increasingly difficult challenge for applied psychology due to the rapid social and technological changes that affect these industries. Deregulation, mergers, increased competition, changes in regulatory practices, and changing public responses are a few of the catchwords that mark the new environmental trends. Applied psychological know-how, especially from organizational psychology and organizational development, is required in at least two areas: analyses of changing work demands and the design of requisite training measures. Applied psychology has in many countries made important contributions to the improvement of system safety via management training and safety-oriented organizational change interventions.

3.3. Safety Culture

Safety culture is a relatively new concept that originated in response to the Chernobyl catastrophe when international analysts and experts were puzzled about the factors that caused the dramatic events of 1986 in the Ukrainian nuclear power plant. The origin of the concept lies in anthropology and organization sciences. Its fundamental assumption is that safety-oriented behavior can only be fully understood by the complex interaction of basic assumptions, values, implicit and explicit norms, and artifacts in the organization. Although a growing number of organizational psychologists have made efforts to obtain a better understanding of the role of organizational cultures, the term safety culture requires further theoretical and methodological refinement. However, the term has gained remarkable currency in many highly hazardous industries far beyond the nuclear industry: aviation, railway safety, chemical

process industry, and medicine. Limited knowledge surrounding the concept of safety culture calls for the decisive involvement of applied psychology.

See Also the Following Articles

Safety and Risks: Errors and Accidents in Different Occupations

Further Reading

Amalberti, A. (2002). Revisiting safety paradigma to meet the safety challenges of ultra complex and safety systems. In B. Wilpert, & B. Fahlbruch (Eds.), *System safety—Challenges and pitfalls of intervention* (pp. 265–276). Amsterdam: Pergamon.

Fahlbruch, B., & Wilpert, B. (1997). Event analysis as problem solving process. In A. Hale, M. Freitag, & B. Wilpert (Eds.), *After the event—From accident analysis to organisational learning* (pp. 113–130). Amsterdam: Elsevier.

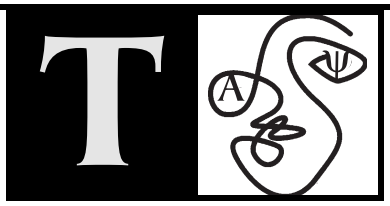
Fahlbruch, B., & Wilpert, B. (1999). System safety—An emerging field of I/O psychology. In C. L. Cooper, & I. T. Robertson (Eds.), *International review of industrial and organizational psychology* (pp. 55–93). Chichester, UK: Wiley.

Graumann, C. G. (1974). Psychology and the world of things. *Journal of Phenomenological Psychology*, *4*, 389–404.

Reason, J. (1997). *Managing the risks of organizational accidents*. Aldershot, UK: Ashgate.

Roland, H. E., & Moriarty, B. (1990). *System safety engineering and management*. New York: Wiley.

Wilpert, B., & Itoigawa, N. (Eds.). (2001). *Safety culture in nuclear power operations*. London: Taylor & Francis.



Teaching Effectiveness

Jac J. W. Andrews
University of Calgary, Calgary, Alberta, Canada

1. Introduction
2. A General View of Teaching
3. Indicators of Effective Teaching
4. Assumptions about Teaching Effectiveness
5. Moving from Effective Teaching to Excellent Teaching
6. Conclusion
   Further Reading

GLOSSARY

direct instruction Teacher-guided instruction aimed at the mastery of basic skills and concepts.

extrinsic motivation The desire to engage in some behaviors to gain outside rewards such as praise, high grades, and money.

intrinsic motivation The desire to engage in some behaviors to gain inside rewards such as satisfaction and insight.

personality trait A stable characteristic of a person that is inferred from a tendency to behave in particular ways in certain situations (e.g., assertive, outgoing, withdrawn).

self-concept A person's awareness of his or her characteristics and the ways in which the person is like and unlike others.

self-efficacy A person's judgment of his or her competence to execute courses of actions that are required to deal with certain situations.

self-esteem The value that a person puts on himself or herself and the value that the person puts on his or her behavior.

strategy Organized skills and processes by which a person plans and regulates his or her behavior and accomplishes particular goals.

Effective teaching is indicated by certain behaviors exhibited in particular situations that lead to positive student outcomes. Essentially, it is when approaches to teaching increase students' knowledge and facilitate their learning. In addition to their content expertise and methodological technique, excellent teachers approach their teaching in ways that are congruent with their students' approach to learning and result in increased student understanding.

1. INTRODUCTION

Although there has been a substantial amount of research over the years in the area of teacher effectiveness, there is still little consensus about how to define and identify effective teaching. One of the reasons for the different opinions about teaching effectiveness is that teaching is such a complex, multidimensional, and idiosyncratic process. However, there are some generally agreed-on notions of teaching and indicators of effective teaching. This article focuses on some of the basic conceptions of teaching and some general understandings of what constitutes effective teaching. From a discussion about teaching and teacher effectiveness, the article concludes with some ideas about moving from effective teaching to excellent teaching.

©2004 Elsevier Inc.
All rights reserved.

2. A GENERAL VIEW OF TEACHING

Teaching is a process aimed at imparting knowledge, facilitating learning, and increasing understanding. Two factors that influence teaching are teacher characteristics and situational–contextual factors (Fig. 1). Teacher characteristics include personality traits, knowledge, abilities, experience, values, and beliefs. Teachers have different types of personality traits, abilities, knowledge, and experiences as well as particular values and beliefs that affect their approaches to teaching. Hence, teachers' approaches to teaching are somewhat dependent on their personal backgrounds, competencies, and viewpoints. Coincidentally, it is generally understood that teachers' competencies vary in relation to their acquired skills and self-efficacy. Moreover, it is generally well known that the values and beliefs of teachers influence their perceptions and judgments and also affect their behavior in the classroom. In 1992, Pratt noted that teachers' values and beliefs are anchored in cultural, social, historical, religious, and personal realms of meaning, and although there are different conceptions about teaching, they are not mutually exclusive. For example, one teacher might have the view that teachers have the primary responsibility for developing students' knowledge. Furthermore, this teacher might believe that students should be rewarded for good work and penalized for lack of effort and poor work. In addition, this teacher might believe that teachers need to direct their students with respect to what they need to learn. Another teacher might believe that the responsibility for learning lies primarily within the learners. This teacher also might believe that students need to know how to learn and understand the importance of their learning and behavior. In contrast to the second teacher, the first teacher would appear to value the product of learning over the process of learning and to consider students as passive learners rather than active learners. In this regard, it is likely that the first teacher would use direct instruction as an approach to teaching and use extrinsic reinforcements to motivate students to learn. The second teacher would likely use an interactive and collaborative approach to teaching and foster an intrinsic motivation for learning within students.

Teachers' approaches to teaching are also mediated by situational and contextual factors. For example, how one teaches will be influenced by things such as the various characteristics of the community

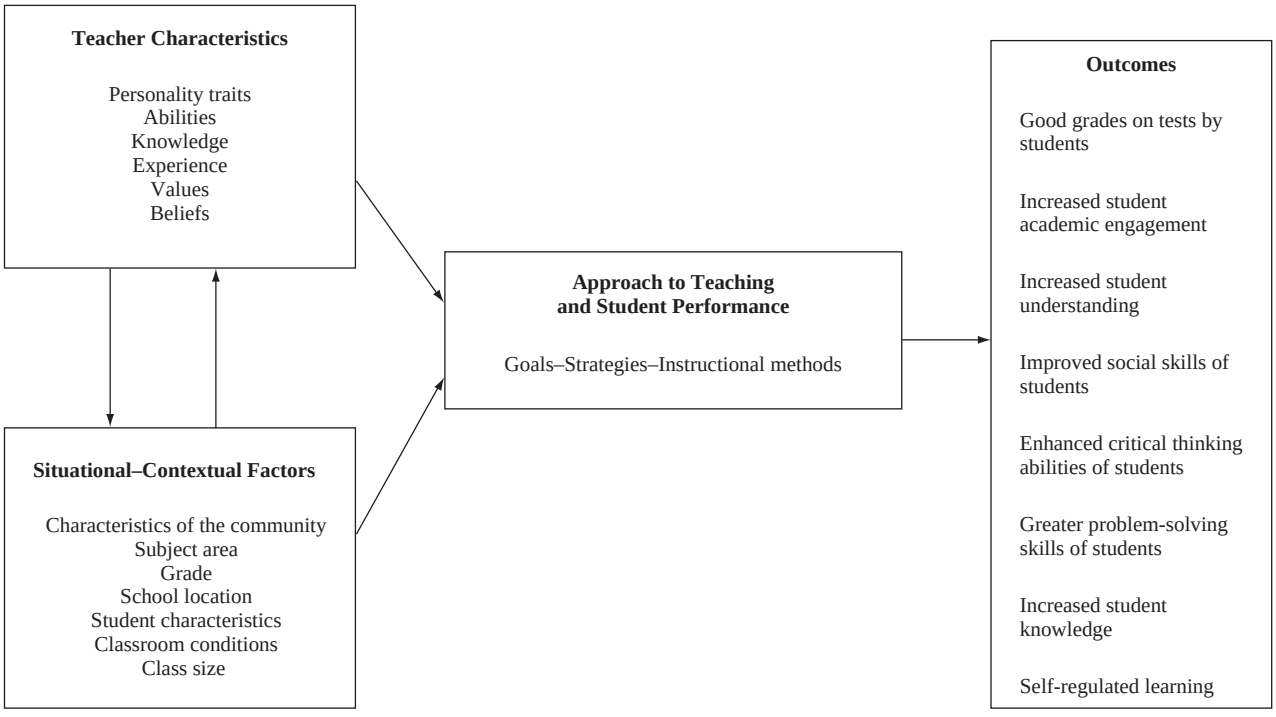


FIGURE 1 General model of teaching.

(e.g., social, political, cultural), the subject area being taught, the grade of the students, the location of the school (e.g., rural or urban), the school district policies, the knowledge and skills of the students, the physical conditions of the classroom, and the class size. For example, teachers will have classrooms that include students with different learning and behavioral styles, abilities, and difficulties that have specific needs and require diverse learning goals, instructional strategies, and methods. Teachers who have students with advanced cognitive abilities must provide opportunities to enrich students' thinking, extend their academic achievement, facilitate their independent learning, and maximize their leadership capabilities. In contrast, teachers who have students with learning disabilities must focus on ways in which to improve their learning strategies, develop their organizational skills, increase their academic self-concepts, and develop their ability to recognize, discriminate, and interpret information. Hence, the various characteristics of students in these classrooms will influence the approach to teaching as well as the outcomes of teaching. For example, in the first case, the general aim is to build on the students' strengths, whereas in the second case, the aim is primarily to remediate the students' weaknesses.

The interaction of teacher characteristics and situational–contextual factors affects the goals, strategies, and instructional methods adopted by teachers in their approaches to teaching. Although there are a number of teaching approaches, the literature indicates that teachers have two primary orientations to teaching: an orientation toward transmitting knowledge and an orientation toward facilitating learning and increasing understanding. Accordingly, each orientation affects the teaching strategies and instructional methods used by a teacher and influences how students approach their learning. For example, if the teacher's goal of teaching is to transmit and increase knowledge, he or she will likely focus on the content (i.e., basic concepts and facts), provide structured learning experiences, and present information in ways that could be reproduced through rote learning. On the other hand, if the teacher wants to maximize students' understanding, he or she will likely focus on the integration and elaboration of information in ways that could facilitate meaningful learning and critical thinking.

Typically, the outcomes of teaching are primarily related to student performance and are the result of teacher characteristics, situational–contextual factors, and the teaching approach. Some of these outcomes could include good grades on tests, increased academic engagement, improved social skills, enhanced critical thinking abilities, greater problem-solving skills, and increased knowledge.

3. INDICATORS OF EFFECTIVE TEACHING

3.1. Teacher Characteristics

In light of the preceding framework of teaching, effective teachers tend to be typified as enthusiastic, charismatic, caring, motivational, fair-minded, supportive, flexible, outgoing, accessible, organized, and adaptable. In terms of abilities and skills, effective teachers tend to prepare for teaching by establishing goals and objectives for their students, determining students' learning strengths and weaknesses, deciding on appropriate instructional strategies and methods, and planning for assessment and evaluation of student performance. During teaching, effective teachers match appropriate learning tasks with students' abilities, execute systematic instructional procedures routinely (e.g., review previously learned concepts and skills, show students the skills they need to be able to perform successfully, explain to students how they can apply these skills to their learning tasks, monitor students' progress, provide feedback to students regarding their learning and development), provide sufficient time for students to engage themselves in their learning, help students to generalize their learning to other situations, communicate with students in ways that promote positive self-concepts, and emphasize student effort and self-control. After teaching, good teachers assess the effectiveness of their teaching strategies and methods, determine the extent to which their goals and objectives have been met, and reevaluate the learning needs of their students.

A notable characteristic of good teachers is their ability to use effective classroom management techniques. For example, effective teachers introduce content and procedures gradually, provide enjoyable and meaningful activities in which students can engage, establish warm and secure classroom environments, use reasonable work standards, ensure high rates of success in assigned activities, anticipate and deal well with problems (e.g., late arrival of students to the classroom), collectively develop rules and routines with the students, promote the social competence of students, stop disruptive behavior appropriately and

immediately, and organize and arrange a classroom that is conducive to student cooperation, active learning, and communication.

Effective teachers tend to have diverse and experience-based knowledge. There is some evidence to suggest that effective teachers have command of the subject area being taught and have a good understanding of students' learning capabilities. Concomitantly, they know how to build on students' background experiences and facilitate their academic and social learning.

3.2. Situational–Contextual Factors

Teaching effectiveness is determined in part by teachers' ability to deal with contextual demands. For example, a good teacher is able to modify his or her teaching approach to adapt successfully to the needs of the students and the requirements of different situations (e.g., teaching in elementary school vs teaching in high school, teaching undergraduate courses versus teaching graduate courses at a university). To some degree, a requisite of effective teaching is being able to determine the situations in which one teaches most effectively. There is some evidence in the literature that teaching effectiveness is situation specific; that is, effective teachers are not necessarily good at teaching all subject areas. Moreover, how effective a teacher is viewed by students may be somewhat dependent on how personable the teacher is and on how meaningful and valuable the students consider the subject area that is being taught. Hence, it is important to keep in mind that because the aim of teaching is to facilitate learning and understanding, the learners' perspective must also be included in any attempt to fully appreciate the nature and scope of an effective teaching–learning transaction.

3.3. Approach to Teaching and Student Performance

There is support in the literature for the view that the quality of students' learning is significantly affected by teachers' approach to teaching. In general, effective teachers tend to take an approach to teaching that encourages students' active engagement in learning, fosters self-regulated learning, results in both increased competency and understanding, and makes learning meaningful, interesting, and enjoyable. Teaching effectiveness is about spending time interacting with students, engaging them in meaningful tasks, and developing their ability to make their own learning choices. It is also about knowing and addressing the individual needs of students and having high expectations for their learning.

4. ASSUMPTIONS ABOUT TEACHING EFFECTIVENESS

In consideration of the preceding discussion and the expansive literature on teaching effectiveness, the following general assumptions are made:

- Although the literature points out a number of characteristics of effective teaching, it is likely that not all effective teachers perform at the highest level in all of these areas.
- Effective teachers are likely able to excel in their teaching without using all of the highly regarded effective instructional strategies and methods.
- Each teacher can be effective in different ways, using different teaching approaches or a combination of teaching approaches with different instructional strategies and methods.
- The outcomes achieved by effective teachers (e.g., students' increased knowledge and grades on tests) may vary from time to time due to situational–contextual factors.
- Knowledge of factors associated with effective teaching is not enough. Teachers must be well trained and mentored.

5. MOVING FROM EFFECTIVE TEACHING TO EXCELLENT TEACHING

Excellent teaching involves processes and practices beyond the descriptors typically associated with effective teaching (e.g., enthusiasm, availability, organization, knowledge of subject matter). It involves more than just content expertise and effective technical performance. Excellent teaching is about focusing on high-quality outcomes (e.g., increased student understanding, self-regulated learning, independent thinking) and ensuring that teaching approaches and students' learning approaches are congruent in achieving these outcomes. In this regard, how students approach their learning is as much a concern for excellent teachers as are the situational–contextual factors and approaches to teaching.

In 1996, Andrews and colleagues discovered that excellent teachers appear to have a conceptual basis for their teaching and value their teaching as well as their students and subject matter. They tend to adopt a meaningful approach to teaching so as to achieve high-quality learning outcomes. In this regard, excellent teachers seem to measure their success in relation to how well their students engage in critical thinking and take responsibility for their own learning. Excellent teachers appear to select teaching approaches that fit with their personalities, abilities, values, and beliefs as well as situational–contextual factors such as subject area and student characteristics.

Excellence in teaching is multifaceted and difficult to achieve. The literature suggests that excellent teachers monitor not only students' misunderstandings but also the incongruencies between teacher and student perceptions of the teaching–learning transaction and desired outcomes. Excellent teaching takes a holistic perspective of the teaching–learning transaction and ensures an integration of process and outcome. It results when there is congruence between teachers' approach to teaching and students' approach to learning.

6. CONCLUSION

Teaching is an instructional process that involves the implementation of strategies and methods relative to particular situations and contexts that are designed to get learners to achieve particular outcomes. Effective teaching is about engaging in particular behaviors that have been shown to lead to desirable student learning outcomes and adapting these behaviors to accommodate diverse learners and educational contexts. Excellent teaching is more than just having effective teaching characteristics and technical performance and being responsive to individual and contextual factors. It is linking students' ways of learning with teachers' ways teaching so that learning becomes authentic, meaningful, and enriching.

See Also the Following Articles

Educational Achievement and Culture ■ Intrinsic Motivation and Self-Determination ■ Learning ■ Literacy, Improvement of ■ Organizational Justice ■ Reading, Teaching of ■ Traits ■ Writing, Teaching of

Further Reading

Andrews, J., Garrison, D. R., & Magnusson, K. (1996). The teaching and learning transaction in higher education: A study of excellent professors and their students. *Teaching in Higher Education*, *1*(1), 81–103.

Beishuizen, J. J., Hof, E., van Putten, C. M., Bouwmeester, S., & Asscher, J. J. (2001). Students and teachers' cognitions about good teachers. *British Journal of Educational Psychology*, *71*, 185–201.

Good, T. L., & Brophy, J. E. (1994). *Looking in classrooms*. New York: HarperCollins.

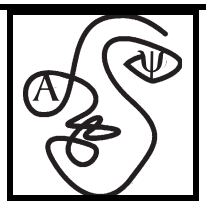
Lortie, D. C. (1975). *Schoolteacher*. Chicago: University of Chicago Press.

Minor, L. C., Onwuegbuzie, A. J., Witcher, A. E., & James, T. L. (2002). Preservice teachers' educational beliefs and their perceptions of characteristics of effective teachers. *Journal of Educational Research*, *96*, 116–127.

Pratt, D. (1992). Conception of teaching. *Adult Education Quarterly*, *42*, 203–220.

Ramsden, P. (1992). *Learning to teach in higher education*. London: Routledge.

Young, S., & Shaw, D. G. (1999). Profiles of effective college and university teachers. *Journal of Higher Education*, *70*, 670–686.



Territoriality

Robert Sommer
University of California, Davis, California, USA

1. The Psychological Concept of Territoriality
2. Methods of Measurement
3. Research Findings
4. Applications
Further Reading

GLOSSARY

defensible space An application of territoriality to public housing that creates areas under the control of residents.
home range An area habitually traversed by an individual or group but not personalized or defended. Home range may be a network of familiar paths and destinations.
jurisdiction A space temporarily controlled by an individual or group.
territory A fixed geographical space marked, defended, and used for life-sustaining activities. At the psychological level, human territory is linked to concepts of personalization and privacy. Turf is a colloquial synonym for territory. The concept has a broader meaning in political science, sociology, and geography, in which it involves national boundary and migration issues.

Territories are areas marked and defended by their owners and used for important life-sustaining activities. Although the concept originated in ethology, there are many differences between human and animal territories. Research on human territory is conducted in real-world settings rather than in the laboratory. The most fruitful applications have been in design, urban planning, and crime prevention.

1. THE PSYCHOLOGICAL CONCEPT OF TERRITORIALITY

Most animals in the wild have fixed geographic areas called territories that they mark, defend against intruders, and use for vital life-sustaining activities. Marking by animals is done in various ways. Some deer species use a glandular secretion rubbed against trees, bears often make claw marks on trees, and birds sing. Humans also mark or personalize and defend spaces important to them—the two major aspects of territory (Fig. 1). However, because humans rely on symbolic means of marking and employ outside agencies to defend their spaces, many differences exist between human and animal territories. Human territories may be less rooted in survival needs than in a desire for status, privacy, and solitude. There may be a complex system of ownership and control in human territories that includes a distant owner, a mortgage holder, a local manager, and tenants. People hire others to patrol and defend their territories (Fig. 2). If the possession is transient, such as when someone places a coat over a chair at a restaurant or a teacher controls a classroom for an hour each day, it seems more useful to describe this temporary occupation as a jurisdiction than as a territory (Fig. 3).

©2004 Elsevier Inc.
All rights reserved.



FIGURE 1 Graffiti announces territorial control of a beach.

In Altman's classification system, primary territories are central to the day-to-day lives of the occupants, clearly marked as theirs, and controlled by them on a regular basis. Examples are a person's home and a private office. Secondary territories are less central and exclusive in a person's life. A neighborhood coffee shop may be a secondary territory because the space, although associated with a specific neighborhood and clientele, is not marked as belonging to a single group and is open to others. Public territories such as sidewalks, streets, and parks, although owned, marked, and policed by a municipality or agency, are open to all those who obey rules of occupancy that may limit usage to certain times and activities.

Territories contribute to a stable social order, sharpen survival skills as occupants patrol and defend them, increase security because people on their turf know its respurces and dangers, and define individual and group identity. Territories increase predictability because people know where to find others, although this may be becoming less true in an era of aspatial and asynchronous digital communication. Because the concept of territory originated in ethology, some of the most promising theories to explain why humans mark and defend territories employ a Darwinian (functional) framework, in terms of the role of territories in natural selection. There are also social learning theories that emphasize its role in the social order and place attachment.



FIGURE 2 Physical barriers used to keep out intruders.



FIGURE 3 Transient occupations are better described as jurisdictions than territories.

2. METHODS OF MEASUREMENT

Although there have been laboratory experiments with animal territories, it is difficult to conduct them with humans. Most studies of human territories are done in the field using natural observation, surveys, and field

experiments. Invading someone's territory as part of an experiment can be stressful and raises ethical issues.

3. RESEARCH FINDINGS

Places with strong territorial definition, such as well-maintained dwellings, tended yards, and many signs of occupancy are less likely to be burglarized and vandalized. Clear territorial definition and many signs of personalization deter criminals, who prefer indistinct border areas where they can remain anonymous.

The more strongly a territory is valued, the more actively it will be defended. Neighbors play an important role in territorial defense.

Dominant individuals in a group or organization are likely to have larger and more suitable territories. Assigned spaces typically reflect the existing dominance order, with higher status individuals receiving the most desirable locations.

People identify themselves with territories, and any threat to the territory becomes a threat to the self. Urban renewal projects are often viewed by local residents as the destruction of a community and a diminution of self.

People on their turf are more opinionated, assertive, and in control. They are more likely to get their way in an argument on their own turf. When children are off their territories, they tend to remain close to a parent. As they become more familiar with a setting, distance between them and the parent increases. Also, males tend to claim larger territories than do women.

Microterritories can be created in small spaces, such as the division of a shared office or dormitory room into two separate territories with each person's belongings on separate sides of the room. The belongings become the territorial markers reinforcing each individual's ownership of half the space. Such microterritories are found in institutional settings in which individuals or groups take possession of certain spaces and keep out intruders. Staff may reinforce such microterritories in order to maintain a stable social order.

4. APPLICATIONS

4.1. Defensible Space

This concept was developed by city planner Oscar Newman to define how public and semipublic areas in public housing can be brought under the control of residents. The two major components of defensible space are surveillance and territoriality. Good surveillance is achieved by proper positioning of windows and video cameras. Residents' feelings of territoriality are enhanced by real and symbolic barriers that mark areas as belonging to particular buildings. Widely used in public housing, defensible space is credited with reducing crime, increasing neighborhood cohesion, and raising property values. This concept is the basis of Crime Prevention through Environmental Design.

4.2. Home Field Advantages

Sports teams play better on their home fields than away. This has been found in football, baseball, hockey, basketball, and soccer. Team members are more familiar with the home field and its idiosyncrasies; are adapted to local weather, lighting, and wind conditions; have local fans cheering for them; can live at home; and are less debilitated by travel. A team will work and play hard an entire season to gain home field advantage for playoff games. Sports owners encourage place identification on the part of both team members and fans, defining the issue in terms of team spirit and fan loyalty.

4.3. Tracking Gang Activities

Urban gangs explicitly mark their turf boundaries using graffiti showing gang symbols. Police departments have officers trained to read wall writing to know what is happening in neighborhoods and head off potential conflict. When symbols from one gang appear on the turf of another gang, authorities may intervene quickly to head off gang warfare. "187" is graffiti code for homicide, a particularly ominous display. Prison authorities are cognizant of informal territories, and new inmates may be cautioned by fellow inmates or staff to avoid sections of the yard or dining hall controlled by gangs. Turf disputes among drug sellers may provoke multiple incidents of violence because of the large amounts of money involved, thereby setting off a cycle of retribution and further violence. Defining the turf boundaries of criminal gangs helps police prevent conflict and make arrests when it occurs.

4.4. Dispute Resolution

Even without established gangs in an area, informal groups may establish spatial boundaries and defend them against intruders. Portions of a school cafeteria may become the de facto territories of cliques and are

off-limits to outsiders. In coastal communities, good fishing or surfing spots may be considered territories by local residents. Graffiti marks a beach for "Locals only." Often described as localism, this attitude can provoke individual or group violence when a stranger intrudes. Sometimes the authorities will formally divide contested space, such as by establishing separate areas for different activities. Segregation is also used to reduce conflict in wilderness recreation, where the main conflict is between motorized and pedestrian visitors. Separate winter trails are established for snowmobiles and cross-country skiers and separate summer trails for hikers and off-road vehicles.

4.5. Urban Planning

It is important to respect territorial boundaries in locating city services. Clients should not be expected to cross through hostile neighborhoods to reach social agencies. A park or playground built between territories of rival gangs is likely to be underutilized and vandalized. Territoriality is included in the "minor incivilities" and "broken windows" theories of community development. Residents are less likely to maintain places that are deteriorated, litter strewn, and over which they lack control. Pride in residency and ownership are concomitants of territoriality.

Design can encourage territories, such as by closing streets in a community or using different color groupings or plantings in a housing development. This gives residents a greater sense of identity and control. Design can also be used to discourage territories in public space, such as a library, a cafeteria, or bus, in which a limited area must be shared and management does not want personal belongings spilling over into adjacent seats.

See Also the Following Articles

Environmental Stress ■ Personal Space ■ Privacy ■ Residential Satisfaction and Perceived Urban Quality

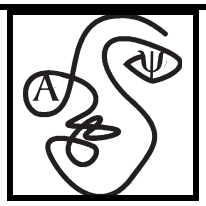
Further Reading

Altman, I. (1975). *The environment and social behavior*. Monterey, CA: Brooks/Cole.

Brown, B. B. (1987). Territoriality. In D. Stokols, & I. Altman (Eds.), *Handbook of environmental psychology* (pp. 505–531). New York: Wiley.

Gifford, R. (2002). *Environmental psychology* (3rd ed.). Colville, WA: Optimal Press.

Taylor, R. B. (1988). *Human territorial functioning*. New York: Cambridge University Press.



Test Anxiety

Moshe Zeidner
University of Haifa, Haifa, Israel

1. Basic and Conceptual Issues
2. Measurement and Assessment of Test Anxiety
3. Anxiety and Cognitive Performance
4. Coping, Interventions, and Clinical Considerations
Further Reading

GLOSSARY

cognitive-focused interventions Emphasize the mediating role of cognitive processes in sustaining or eliminating test anxiety and refer to a wide array of therapeutic approaches directed toward modifying the worry and irrational thought patterns of test-anxious clients.

debilitating anxiety Refers to anxiety that is predominantly harmful to task performance.

emotionality Consists of perceptions of autonomic reactions evoked by evaluative stress.

emotion-focused interventions Aim primarily at reducing the arousal and heightened emotional reactions of test-anxious people when faced with stressful evaluative situations.

situation-specific personality trait In the context of test anxiety research, refers to the individual's disposition to react with extensive worry, intrusive thoughts, mental disorganization, tension, and physiological arousal when exposed to evaluative situations.

test anxiety The set of phenomenological, physiological, and behavioral responses that accompany concern about possible negative consequences or loss of competence on an exam or similar evaluative situation. Test-anxious behavior is typically evoked when a person believes that his or her intellectual, motivational, and social capabilities are taxed or exceeded by demands stemming from the test situation.

worry Refers primarily to cognitive concern about the consequences of failure.

This article examines test anxiety theory, measurement, performance, and clinical parameters. First, it provides a historical backdrop for test anxiety research and presents a number of conceptualizations and distinctions, delineating key facets and components of test anxiety. Next, it discusses measurement and assessment issues, focusing on self-report and alternative assessment procedures. Then, the anxiety–performance interface is discussed, noting key moderating and mediating factors. It concludes by presenting a number of treatment and intervention programs and clinical considerations when treating test-anxious clients.

1. BASIC AND CONCEPTUAL ISSUES

1.1. Overview

Tests and evaluative situations have emerged as a potent class of anxiety-evoking stimuli in our society, which bases many important decisions relating to an individual's status in school, college, and work on tests and other assessment devices. Thus, test anxiety figures prominently in the literature as one of the key villains in the ongoing drama surrounding psychological testing and is frequently cited among the factors at play in

©2004 Elsevier Inc.
All rights reserved.

determining a wide array of unfavorable outcomes and contingencies, including poor cognitive performance, scholastic underachievement, and psychological distress and ill health. Many individuals have the ability to do well on exams but perform poorly because of their debilitating levels of anxiety. Consequently, test anxiety may limit educational or vocational development because test scores and grades influence entrance to many educational or vocational training programs in modern society. The field of test anxiety research has prospered, in part, due to the increasing personal salience of test situations for people in modern society, making tests and their long-term consequences significant educational, social, and clinical problems for many.

Experience of test anxiety is a near-universal phenomenon across people differing in age, gender, and culture. Meta-analyses of test anxiety data from various national sites show that although mean test anxiety levels vary somewhat across cultures, test anxiety is a prevalent and relatively homogeneous cross-cultural phenomena. Furthermore, women tend to report higher levels of test anxiety than men, but the gender difference often does not translate into objective performance differences. Gender differences are attributable to differential exposure and learning experiences as well as the possibility that women are more willing to admit they are anxious.

This article first provides an overview of basic issues and conceptualizations. Then, it discusses the measurement and assessment of test anxiety, followed by a review of empirical studies of evaluative anxiety and performance. Finally, it examines interventions tailored to alleviate test anxiety and its affective, cognitive, and behavioral components.

1.2. Conceptions of Test Anxiety

The term test anxiety refers to the set of phenomenological, physiological, and behavioral responses that accompany concern about possible negative consequences or loss of competence on an exam or similar evaluative situation. Test-anxious behavior is typically evoked when a person believes that his or her intellectual, motivational, and social capabilities are taxed or exceeded by demands stemming from the test situation.

Throughout its relatively brief history as a scientific construct, test anxiety has taken on a variety of different meanings as a function of both changes in the "zeitgeist" and variations in the theoretical persuasions of the individual investigators involved in this research arena. Thus, in the early days of test anxiety research the construct was initially defined in motivational terms as drive level, goal interruption, or need to avoid failure. Test anxiety was also conceptualized as a relatively stable personality disposition that develops when parents hold exaggerated expectations and are overcritical of their children's achievement efforts. In keeping with the cognitive revolution of the late 1960s and early 1970s, test anxiety came to be viewed primarily as a cognitive–attentional phenomenon, as evidenced by a number of key chapters in Sarason's 1980 book on test anxiety. Accordingly, the highly anxious person is one who attends, in an exaggerated way, to evaluative cues, to self-generated concern about the ability to do well enough, and to feelings of physiological arousal. Recent conceptualizations have viewed test anxiety as part of a cybernetic self-control process, with anxiety reflecting the existence of a conflict between competing reference values, or as a form of self-handicapping employed to preserve one's self-merit in the face of potential failure.

The test anxiety construct was dramatically advanced by a number of important conceptual distinctions, which helped refine thinking and research in the area. One useful distinction, advanced by Charles Spielberger, is between anxiety as a relatively stable personality trait and anxiety as a more transitory state reaction to specific ego-threatening situations. Another important conceptual and methodological contribution to the test anxiety literature is the distinction made by Alpert and Haber between facilitating and debilitating anxiety. Accordingly, facilitating and debilitating anxiety are claimed to lead to task-related and task-irrelevant behaviors during test situations, respectively. Yet another conceptual contribution, advanced by Liebert and Morris, is the critical differentiation between worry and emotionality components of test anxiety. This distinction proved to be instrumental in shifting test anxiety theory and research toward a more cognitive orientation. According to this conceptualization, anxiety was viewed to be a bidimensional phenomenon, including cognitive (worry) and affective (emotionality) components. Specifically, worry was viewed primarily as cognitive concern about the consequences of failure, whereas emotionality was defined as consisting of perceptions of autonomic reactions evoked by evaluative stress. These two components are revealed to be empirically distinct, though correlated, and worry relates more strongly to test performance than does emotionality.

Theorizing by Zeidner, in his 1998 book on test anxiety, emphasized the distinction between test

anxiety as an attribute of the person and as a dynamic process. From the first perspective, dispositional test anxiety may be construed as a contextualized personality trait. Accordingly, test anxiety refers to the individual's disposition to react with extensive worry, intrusive thoughts, mental disorganization, tension, and physiological arousal when exposed to evaluative contexts or situations. The more transient state expressions of anxiety may be assessed separately from the more stable trait. From the second, process-oriented perspective, test anxiety depends on the reciprocal interaction of a number of distinct elements at play in the ongoing stressful encounter between a person and an evaluative situation, including the evaluative context, individual differences in vulnerability (trait anxiety), threat perceptions, appraisals and reappraisals, state anxiety, coping patterns, and adaptive outcomes. Events that elicit test anxiety consist of a number of distinct temporal phases, including preparation, confrontation, anticipation, and resolution. Accordingly, threat appraisals, state anxiety levels, and competence for task performance may change at different stages.

Much of the ambiguity and semantic confusion associated with the status of test anxiety as a psychological construct stems from the fact that different investigators have studied this term with quite divergent meanings. Thus, test anxiety has been used to refer to several related but logically very different constructs, including stressful evaluative stimuli and contexts, individual differences in anxiety proneness in evaluative situations (i.e., trait anxiety), and fluctuating anxiety states experienced in a test situation (i.e., state anxiety). Although the question still looms large whether test anxiety is best conceptualized as a relatively stable personality trait (individual difference variable) or an ephemeral emotional state, a widely accepted conceptualization by Spielberger proposes a hybrid term by construing test anxiety as a situation-specific personality trait. Accordingly, test anxiety refers to the individual's disposition to react with extensive worry, intrusive thoughts, mental disorganization, tension, and physiological arousal when exposed to evaluative situations.

1.3. Facets and Typologies

In contrast to early mechanistic views of test anxiety as a unified construct, test anxiety is currently construed as a complex multidimensional construct embodying a series of interrelated cognitive, affective, and behavioral components and reactions. The fact that anxiety is such a complex construct, encompassing worry and self-preoccupation, physical upset, disruptive feelings, and maladaptive behaviors, makes it particularly difficult for researchers to sort out all these components. Although a number of attempts have been made to determine the dimensionality of the test anxiety construct and to tease out its key facets, little agreement has been reached among experts on the exact number of facets or components of test anxiety. However, researchers have found it particularly useful to differentiate between a cognitive facet (worry, irrelevant thinking, etc.), an affective facet (tension, bodily reaction, and perceived arousal), and a behavioral facet (deficient study skills, procrastination, avoidance behaviors, etc.). Thus, test-anxious individuals may be characterized by their thoughts, somatic reactions, feelings, and observable behaviors in evaluative situations. In any given test situation, test-anxious subjects may experience all, some, or none of these test anxiety reactions. The specific anxiety response manifested may vary, depending on the constitutional qualities and past experience of the individual, the nature of the problem to be solved, and various situational factors affecting the level of anxiety evoked.

Discussions of test anxiety in the literature are commonly guilty of a "uniformity myth," conveying the impression that test anxiety is a rather homogeneous or unidimensional category. However, under the assumption that test anxiety is a multidimensional and multidetermined phenomenon, it stands to reason that a variety of different types of test-anxious examinees may be identified. This simple fact is often overlooked when writers present theory and research relating to the "test-anxious" student—typically treated as a uniform category. Some examinees may be anxious in test situations because they have high motivation to succeed on academic tasks given the likelihood of failure, some may have poor study or test-taking skills, some may be anxious because they have low intellectual ability, some tend to be perfectionist overachievers and will be dissatisfied with anything less than a perfect score, and others are anxious because they fail to meet social expectations or fear parental punishment.

2. MEASUREMENT AND ASSESSMENT OF TEST ANXIETY

2.1. Self-Report Instruments

Self-report inventories have become the most prevalent format for assessing test anxiety because they are

considered to provide the most direct access to a person's subjective experiential states in evaluative situations; possess good psychometric properties; are relatively inexpensive to produce; and are simple to administer, score, and interpret.

Among the most widespread contemporary measures of test anxiety are Spielberger's Test Anxiety Inventory, Sarason's Reactions to Tests, Suinn's Test Anxiety Behavior Scale, and Benson *et al.*'s Revised Test Anxiety Scale.

A test anxiety scale needs to show adequate reliability and validity before it can be used in making practical decisions. Fortunately, most popular test anxiety inventories have satisfactory reliability coefficients, typically in the high .80s to low .90s. During longer intervals between assessments, personality traits such as test anxiety may change, causing lower stability coefficients. Additional factors influencing reliability are test length, test–retest interval, variability of scores, and variation within test situation.

Factor analysis, both exploratory and confirmatory, has been the method of choice in studying the structure and construct validity underlying items on test anxiety inventories. A large number of factor analytic studies have supported two to four factor structures of the construct. However, self-report scales are often plagued by a number of conventional threats to validity, including (i) response bias—that is, the tendency of a person to favor a particular response regardless of the stimulus (e.g., respond "yes" to the item "I feel calm and relaxed during important exams"); (ii) social desirability—that is, responding in ways judged socially acceptable to the self, examiner, or others expected to see the results; (iii) defensiveness—persons who are high in defensiveness may be experiencing test anxiety at some level but may be repressing conscious feelings about it; and (iv) fakeability—students may be able to fake bad (high test anxiety levels) or fake good (low test anxiety levels) on these measures.

A main problem confronting criterion validity lies in finding an agreed upon or acceptable criterion; currently, there is no infallible or perfectly objective criterion against which to validate test anxiety scores. Scores on ability tests, grade point average, observer ratings, behavior in structured evaluative situations, and the like have been employed as measures of criterion behaviors. It is noted that if anxiety is expected to be evoked in the criterion performance, then attempts to minimize anxiety in a selection instrument may in fact reduce its validity as a predictor.

2.2. Alternative Assessment Procedures

A variety of less frequently used assessments have been employed in assessing test anxiety, including "think-aloud" procedures (e.g., "Please list as many thoughts and feelings as you can recall having during this test"), physiological measures designed to gauge changes in somatic activity believed to accompany the phenomenological and behavioral components of test anxiety, trace measures (e.g., accretion levels of corticosteroids, adrenaline products, sugar, cholesterol, and free fatty acids), performance measures (e.g., examination scores, semester grade point averages, and latency and errors in recall of stress-relevant stimulus materials), and systematic observations of specific behaviors reflective of test anxiety (perspiration, excessive body movement, chewing on nails or pencil, hand wringing, "fidgety" trunk movements, and inappropriate laughter when subjects were engaged in exam situations).

2.3. Practical Uses in Applied Settings

The ultimate test for a test anxiety measure is that it shows practical utility in applied settings, and current measures leave much to be desired in this respect. Seldom do test anxiety scales inform us about the various situational and personal factors eliciting test anxiety (e.g., anxiety proneness, inadequate preparation, and overstimulation), the full range of manifestations of test anxiety (e.g., cognitive, affective, and behavioral), coping procedures and strategies, the consequences of test anxiety, or the dynamic fluctuations in test anxiety states across various phases of a stressful evaluative encounter. Future scales need to be more relevant for planning, execution, and evaluation of clinical or educational interventions through specification of the various antecedent conditions, manifestations, and consequences of test anxiety. Furthermore, current instruments are designed to measure the relative presence of test anxiety but do not inform us enough about the very high or very low test-anxious individual, which is essential for clinical decision making. Thus, future measures need to cover specific manifestations of anxiety, ranging from a total lack of concern about evaluation and minimal motivation to debilitating levels of worry, anxiety blockage, and panic. In addition, future measures need to distinguish between facilitating and debilitating arousal and cognitive processes that are realistic (e.g., worrying about an imminent difficult exam in physics, which is prompted

by a genuine threat) and those that are unrealistic (e.g., those prompted by an unlikely exam failure).

3. ANXIETY AND COGNITIVE PERFORMANCE

The importance of test anxiety as a key construct in understanding sources of examinee distress, impaired test performance in evaluative situations, and academic underachievement is readily apparent. This situation demands that test anxiety be better understood through systematic assessment and research and appropriately dealt with. Indeed, much of test anxiety research during the past half century has been conducted to help shed light on the negative effects of test anxiety on examinee performance and these concerns have stimulated the development of a variety of assessment methods as well as therapeutic techniques and intervention programs.

3.1. Nature of Anxiety–Performance Relationship

Hundreds of studies have investigated the complex pattern of relations between anxiety and different kinds of performance. Test anxiety has been found to interfere with competence both in laboratory settings and in true-to-life testing situations in school or collegiate settings. Processing deficits that relate to test anxiety include general impairments of attention and working memory, together with more subtle performance changes such as failure to organize semantic information effectively.

Hembree's 1998 meta-analytic study, based on 562 North American studies, demonstrated that test anxiety correlated negatively, although modestly, with a wide array of conventional measures of school achievement and ability at both high school and college levels. Data collected on students from the upper elementary school level through high school show that test anxiety scores were significantly related to grades in various subjects, although the correlation was typically approximately –0.2. Cognitive measures (i.e., aptitude and achievement measures combined) correlated more strongly with the worry than emotionality component of test anxiety. Higher effects sizes were reported for low- than high-ability students and for tasks perceived as difficult than those perceived as being easy. Another meta-analysis reported by Ackerman and Heggestad in 1997 showed a mean correlation of –.33 between test anxiety and general intelligence test performance. Test anxiety was also correlated in the –.20 to –.30 range with other broad intellectual abilities, including fluid and crystallized intelligence, learning and memory, visual perception, and math ability.

The nature of the anxiety–competence relationship is best viewed as reciprocal in nature. Thus, high levels of test anxiety, with elevated levels of worry and cognitive interferences, absorb part of the capacity needed for attention, short-term memory, problem-solving, or other cognitive processes required for successful completion of the test. Test anxiety also produces certain aversive patterns of motivation, coping, and task strategies that interfere with learning and performance. The result is that competence and self-efficacy suffer, thus leading to further anxiety over time and generating a vicious circle of increasing anxiety and degrading competence. Future research would profit from employing nonrecursive process models in order to better capture the dynamic and cyclical nature of the anxiety–performance relationship.

3.2. Mediating and Moderating Effects

3.2.1. Mediating Effects

There is a large literature on test anxiety as a predictor of information processing in laboratory studies that overlaps with studies of general anxiety. The information-processing components sensitive to test anxiety relate to input (encoding and acquisition of information), central processing (e.g., memory, language processing, conceptual organization, judgment, and decision making), and output (e.g., information retrieval, response selection, and execution). Deficits related to test anxiety have been identified at various stages of processing, suggesting some general impairment in attention and/or working memory. These various performance deficits are often attributed to high levels of worry and cognitive interference.

Both cognitive interference and cognitive bias appear to be pervasive in evaluative anxiety, influencing various stages of information processing. In general, these mechanisms appear to operate much as they do in general anxiety, although evidence is lacking on cognitive bias and evaluative anxiety. A hypervigilance theory put forth by Michael Eysenck in 1992 plausibly suggests that anxiety leads to scanning of the environment for threat (generating distractibility and attentional impairment), followed by focusing of attention on sources of

threat (generating attentional bias). In addition, competence deficits may also be a consequence of poor skill acquisition. For example, deleterious effects of test anxiety may reflect not just cognitive interference but also deficits in study habits and test-taking skills.

Behavioral avoidance generated in part by performance-avoidance goals plays a key role in maintenance of evaluative anxiety and concomitant skill degradation. Test anxiety leads to procrastination, motivated by the aversiveness of the test material or fear of failure on the test. Procrastination, such as failure to complete homework assignments or study for the test, leads to failure to acquire the knowledge required. In turn, this lack of preparation leads to poor performance and anxiety in the test situation, increasing subsequent test anxiety and avoidance of study.

3.2.2. Moderating Effects

Studies also identify moderator variables that accentuate or reduce deficits in performance. For example, negative feedback appears to be especially detrimental to test-anxious subjects, whereas providing reassurance and social support may eliminate the deficit. However, there have been sufficient instances of nonconfirmation of predicted deficits to suggest that high anxiety does not automatically generate lower achievement outcomes.

Although anxiety is predominantly harmful to task performance (debilitating anxiety), it may sometimes have a positive effect (facilitating anxiety). One of the factors that may especially tip the scales toward debilitating effects is the presence of worry because of its tendency to produce distracting cognitive interference. The nature of the task may also play an important moderating role. Generally, test anxiety is more detrimental to attentionally demanding tasks and may even facilitate performance on easy tasks. There may also be more subtle effects related to the qualitative nature of the task.

4. COPING, INTERVENTIONS, AND CLINICAL CONSIDERATIONS

4.1. Coping

Text anxiety relates to higher emotion-focused coping (e.g., trying to control anxiety symptoms) and avoidance (e.g., trying not to think of the test) but to lower task-focused coping (e.g., focusing effort on task performance). Emotion-focused coping and avoidance both appear to predict state anxiety in evaluative situations. Perhaps surprisingly, task-focused coping also relates to higher pretest anxiety: Exam preparations may inevitably lead to elevated anxiety. It is often difficult to categorize coping strategies as exclusively adaptive or maladaptive, but task-focused coping tends to lead to higher grades than avoidance, although the data are mixed.

4.2. Intervention Programs and Techniques

A bewildering array of test anxiety treatment programs have been developed and evaluated during the past three decades. Treatment fashions and orientations have swayed sharply from the clinical to the behavioral, and recently to the cognitive perspective—essentially mirroring the evolution of the behavior therapies. There is no simple organizing principle with which to categorize the plethora of therapeutic techniques and approaches that have proliferated during the past few decades. Current attempts to reduce debilitating levels of test anxiety and enhance test performance have typically focused on treatments directed toward the emotional (affective), cognitive (worry), or behavioral skills (study skills) facets of test anxiety. Thus, treatment programs typically include emotion-focused treatments, designed largely to alleviate negative emotional affect experienced by test anxious persons; cognitive-focused treatments, designed to help the test-anxious client cope with worry and task-irrelevant thinking; and study-skills training, designed to improve study and test-taking skills and enhance test performance.

A brief summary of key emotion-focused, cognitive-focused, and skill-focused treatment techniques and methods, and their reported effectiveness, is presented in Table I.

4.2.1. Emotion-Focused Interventions

The emotionally oriented therapies aim primarily at reducing the arousal and heightened emotional reactions of test-anxious persons when faced with stressful evaluative situations. Based on the assumption that anxiety comprises a physiological component, attempts to alleviate anxiety symptoms should prove successful, in part, if they focus on reducing levels of arousal or on altering ways in which people appraise their arousal in evaluative situations.

The basic strategy in these treatments is directed to teach the client certain skills (mainly relaxational) so that when confronted by stress-inducing evaluative

TABLE I

Focal Test Anxiety Intervention Techniques

Treatment	*Description*	*Effectiveness*
Emotion-focused intervention		
Biofeedback	Use of instrumentation (e.g., a physiograph) to provide a person with immediate and continuous information about one or more physiological processes (e.g., skin conductance, temperature, heart rate, blood volume pulse, respiration, and electromyograph). Biofeedback teaches high test-anxious persons to monitor and modify the physiological processes associated with their emotional reactions.	A large body of literature supports the notion of *increased* physiological control when using physiological feedback and self-regulation. However, biofeedback alone is not effective in reducing test anxiety (nor does the addition of biofeedback training improve the efficacy of other forms of treatment). Given the potential cost and inconvenience of using biofeedback training, it may not be the treatment of choice for test anxiety intervention.
Relaxation training	Recommended on the premise that maintaining a relaxed state during testing, via deep breathing and muscle relaxation exercises, would counteract a person's aroused state. Presumably, if a person knows when and how to apply relaxation, it will be applied directly as a counterresponse to anxiety.	Meta-analytic research tends to support the effectiveness of relaxation therapy. However, the effects on performance tend to be negligible.
Systematic desensitization	Test anxiety is viewed as a classically conditioned emotional reaction resulting from a person's aversive experiences in evaluative situations. Systematic desensitization proposes that anxiety reactions to test situations may also be unlearned through specific counterconditioning procedures. The test-anxious client is typically trained in a deep muscle relaxation procedure and, while relaxed, instructed to visualize an ordered series of increasingly stressful test-related scenes (an "anxiety hierarchy"). The client imaginally proceeds up the hierarchy until he or she is able to visualize the most stressful scenes on the list without experiencing anxiety. Through repeated pairings of imaginal representations of threatening evaluative situations with deep relaxation, the bond between the threatening evaluative scenes and anxiety is expected to be weakened.	Meta-analytic data lend support to the effectiveness of systematic desensitization in reducing test anxiety in school-children and college students. It is shown to be as effective, if not more effective, in reducing test anxiety than a variety of other treatments, including relaxation training, hypnosis, and study skills training. However, systematic desensitization fares less well when cognitive performance (e.g., academic achievement) is the criterion or outcome being assessed.
Anxiety management	Teaches highly test anxious subjects to recognize their test-related arousal responses as they are building and then to use them as cues for initiating the coping response of relaxation in evaluation situations.	A body of research supports the effectiveness of this technique in reducing test anxiety. Thus, anxiety management training appears to be as robust and effective, if not more so, than related interventions. Reductions in debilitating test anxiety were maintained for follow-up periods ranging from several weeks to several months.

Continues

Continued

Treatment	*Description*	*Effectiveness*
Modeling	Involves the live or symbolic (e.g., through videotape) demonstration of desired coping behaviors in a stressful evaluative situation such that they can be subsequently imitated by the test-anxious person. It is assumed that exposure to models displaying adaptive behavior may play a positive role in facilitating performance. Clients are instructed to vividly imagine the stressful evaluative scene and to focus on the anxiety and associated response-produced cues (e.g., racing heart, neck and shoulder tensing, dryness of the mouth, and catastrophic thoughts). The clients are then trained to use these cues to prompt adaptive coping skills to actively relax away tension and reduce anxiety before it mounts too severely.	A body of research lends support to the effectiveness of modeling in treating test anxiety. In particular, exposure to models who are task-oriented and provide attention-directing cognitive structuring clues is beneficial to the performance of test-anxious persons. Of additional benefit is evidence in the behavior of the model that he or she is successfully coping with the worry and tension associated with test anxiety.
Cognitive-focused interventions		
Cognitive–Attentional training	Cognitive–attentional training provides specific training in the redirection of attention to task-focused thinking and emphasizes the inhibition of task-irrelevant thinking and nonproductive worry. The cognitive–attentional approach relates performance decrements to the diversion of attention to self-focused thinking, coupled with the cognitive overload caused by the worry component of anxiety. By redirecting attention to the task and reducing worry and task-irrelevant thinking, cognitive resources are freed, and when redirected to the task, performance is improved. Attentional training programs traditionally provide clients with instructions to attend fully to the task and to inhibit self-relevant thinking while working on a variety of academic tasks.	The beneficial effects of attentional instruction on the anxiety and cognitive performance of high test-anxious students is supported by some empirical research. Task instructions that provide examinees with information about appropriate problem-solving strategies, and away from self-preoccupied worry, may be particularly helpful to the test-anxious individual's cognitive functioning.
Cognitive restructuring	The rationale is that test-anxious persons will be able to master their anxiety by learning to control task-irrelevant cognitions that generate their anxiety and direct attention from their task-directed performance. The two most prominent cognitive therapeutic methods in test anxiety intervention are rational emotive therapy and systematic rational restructuring. Both forms of treatment are based on the premise that anxiety or emotional disturbance is a result of illogical or "irrational" thinking. Two key irrational beliefs that maintain test anxiety is that one *must* succeed on major exams at all cost and that success on these exams is equivalent to self-worth.	Research indicates that although cognitive restructuring reduces anxiety, there is no concomitant improvement in performance. A number of studies provide evidence showing that these technique may be effective in reducing anxiety. However, concomitant improvements in cognitive performance are observed with far less consistency.

Treatment	*Description*	*Effectiveness*
	Test-anxious individuals are taught how to recognize, vigorously challenge, question, and dispute their irrational beliefs and replace their maladaptive internal dialogue with more rational structures and beliefs. Presumably, by modifying irrational beliefs and schemas, negative emotional reactions will be reduced and performance improved. Systematic rational restructuring aims at helping test-anxious clients to discover the worrisome task-irrelevant thoughts they entertain during tests, to eclipse such thoughts, and to substitute positive self-statements that redirect their attention to the task at hand.	
Cognitive–behavioral modification	A multifaceted program merging both cognitively focused and emotionally focused techniques (as well as skill training in many cases), thus offering the test-anxious client the best of many worlds, so to speak. This multimodal treatment attempts to deal with the multiple manifestations of test anxiety, including negative motivational or affective tendencies, irrational thought patterns, and skills deficits, and emphasizes the application and transferring of acquired coping skills to *in vivo* test situations. Given its dual emphasis on modifying both emotional processes and irrational thoughts and cognitions, this results in a powerful approach that merges emotionally oriented and cognitively oriented techniques to alleviate clients' test anxiety and enhance their test performance. This procedure is based on the premise that reducing a person's level of test anxiety involves both anxiety reduction training and detailed cognitive restructuring of certain faulty beliefs or misconceptions concerning evaluative situations. It is also assumed that some students may require restructuring of their study habits if rewarding and successful academic work is to be ensured.	"Multimodal" treatment packages, such as cognitive–behavior modification, are most likely to be effective by their support for the inclusion of multiple domains related to test anxiety. These procedures are relatively effective in reducing self-reported levels of debilitating test anxiety and are equally effective, more or less, in reducing both cognitive and affective components of test anxiety test performance, on average, by about half a standard deviation in school-aged samples and elevate grade point average by approximately three-fourths of a standard deviation.
Skill-focused interventions		
Study skills training	Study skills training differs from the other cognitive therapies in that it does not directly focus on modifying the cognitive component of anxiety but rather centers on improving students' study and test-taking skills. Presumably, improvement of these skills should have a direct impact on performance through improved mastery of the test material and also indirectly impact performance through the reduction of worries surrounding inadequate exam preparation.	These programs have been successful mainly in enhancing the test performance of high test-anxious subjects with poor study skills. It is apparent that when test-anxious subjects suffer from serious study skill or test-taking deficits, alternative forms of treatment would not be expected to lead to performance gains, inasmuch as the behavioral deficit still exists. At the same time, skill training alone may lead a person to performance gains, but it may still leave the person distressed.

situations in the future he or she will be able to handle them adequately. The therapies also provide opportunities for application of training either within the therapy setting or in real-life situations.

These emotion-focused procedures typically include a number of common components, such as theoretical explanations of test anxiety as a conditioned response and the "deconditioning" rationale for treatment; instructions in specific methods for reducing anxiety, such as relaxation and guided imagery; guided practice in therapeutic methods; and practice (homework and *in vivo* practice). By and large, these emotion-focused treatments rely on key behavioral learning principles (counterconditioning, reciprocal inhibition, extinction, observational and coping skill learning, etc.) and also draw from an arsenal of behavioral techniques, such as deep muscle relaxation, guided imagery, and graduated hierarchies. For example, relaxation and guided imagery is not unique to a particular test anxiety behavioral intervention method, but it is employed in several methods, including relaxation as self-control, systematic desensitization, and anxiety management training.

Procedures designed to reduce emotionality, although clearly useful in modifying subjectively experienced anxiety, by themselves appear to have little effect on cognitive performance. Overall, emotion-focused treatments appear to be relatively ineffective in reducing test anxiety unless these treatments contain cognitive elements. It may therefore be necessary to combine such approaches with therapy modes focusing specifically on cognitive change in order to reliably elicit improvement in cognitive performance.

4.2.2. Cognitive-Focused Interventions

Recently, there has been a proliferation of cognitively oriented intervention programs that emphasize the mediating role of cognitive processes in sustaining or eliminating test anxiety. In part, the documented failure of emotionally oriented behavioral therapies to markedly improve the academic performance of test-anxious students, coupled with the inconsistent relation reported between emotional arousal and test performance, has led to a greater emphasis on cognitive factors in test anxiety intervention. Indeed, reviews of the literature conclude that cognitively based treatment strategies are more powerful than direct behavioral therapies in effecting test anxiety and performance changes.

Cognitive therapy is a generic term that refers to a wide array of therapeutic approaches directed toward modifying the worry and irrational thought patterns of test-anxious clients. Broadly speaking, cognitively oriented approaches to test anxiety intervention are quite similar in assuming that cognitive processes are determining factors in test anxiety, although they differ in terms of actual intervention procedures. A fundamental assumption shared by contemporary cognitive models of test anxiety is that cognitive processes mediate the person's emotional and behavioral responses to stressful evaluative situations. It follows that in order to modify the negative emotional reactions of test-anxious clients to evaluative situations, therapy needs to be directed at reshaping the faulty premises, assumptions, and negative attitudes underlying maladaptive cognitions of test-anxious subjects.

Given their emphasis on modifying emotional processes, irrational thoughts and cognitions, and behavioral deficits, this results in a powerful approach that merges emotionally oriented, cognitively oriented, and behaviorally oriented techniques to alleviate clients' test anxiety and enhance their test performance.

4.2.3. Skill-Focused Intervention

Recent thinking and research suggest that high test-anxious students with poor study and test-taking skills would probably benefit most from study skills training intervention intended to improve their study and test-taking habits and skills. This intervention is based on a cognitive-deficit or deficiency model. Whereas study skills training is directed toward improving a variety of cognitive activities that affect the organization, processing, and retrieval of information (e.g., study habits and test-taking skills), training in study skills does not directly address the specific cognitive concomitants of test anxiety. Instead, it is designed to augment other cognitive interventions.

The distinction between the various treatment orientations is quite fuzzy, and these approaches are becoming increasingly difficult to distinguish. Although there may be highly specific interventions, which have an affective (e.g., relaxation therapy) or cognitive (e.g., rational emotive therapy) orientation, most methods are normally embedded in a multidimensional context. A combination of procedures (whether combined in a truly integrative manner or in the stance of technical eclecticism) seems to best represent the true nature of the test anxiety intervention process.

4.3. Clinical Considerations

This section presents a number of important considerations that should be kept in mind by both researchers

and practitioners when developing, implementing, or evaluating test anxiety intervention programs.

4.3.1. Performing Careful Diagnosis of Client's Problem

In order to tailor treatment programs to meet the specific needs and problems of the client, a logical first step is a careful diagnostic assessment and analysis of the nature of the test-anxious person's affective and cognitive problem(s). For some test-anxious subjects, provision of skills training may be the treatment of choice, whereas for others it would involve building up of self-confidence in a particular content area (e.g., math) or teaching relaxation skills. Information about the following aspects of the client's problem might be particularly useful: nature of the problem as experienced and defined by the test-anxious client, perceived severity and generality of the problem, duration and extent of test anxiety, perceived origins of test anxiety, situation-specific factors that intensify or alleviate anxiety reactions, specific consequences of test anxiety for the client, and suggested changes that the client views as potentially helpful. A careful diagnostic assessment may suggest factors other than test anxiety that underlie one's heightened emotional reactions in test situations.

4.3.2. Meeting Preconditions for Therapeutic Effectiveness

In order for a test anxiety intervention program to work, a number of preconditions need to be met. First, test-anxious individuals should possess certain relevant skills in their behavioral repertoire (e.g., problem-solving, relaxation, study/test-taking skills) to apply in appropriate evaluative circumstances. Second, test-anxious clients must be sufficiently motivated to deal directly with evaluative situations and have the wherewithal and self-efficacy to efficiently implement the coping skills they have at their disposal. Third, test-anxious persons must be provided with an adequate amount of practice and experience in applying various coping skills in true-to-life evaluative situations in order to ensure transfer of therapy from the treatment environment to the real world.

4.3.3. Adjusting Treatment to the Needs of Particular Types of Test-Anxious Individuals

Interventions and therapeutic techniques would be most effective if they could be adjusted to suit the needs of different types of test-anxious persons. Because there are different types of high test-anxious individuals, as suggested previously, each characterized by different problems and concerns (e.g., failure in meeting personal or social expectations, low feelings of self-efficacy and failure acceptance, and poor study skills), no single treatment program would be expected to be equally effective for all clients. Thus, for some highly perfectionist test-anxious individuals therapy may focus on lowering socially prescribed performance expectations, whereas for other "failure accepting" students therapy may consist of raising performance expectancies and enhancing perceived self-efficacy. Students high in test anxiety with sound study or test-taking skills should profit from treatment focusing on test anxiety reduction. In contrast, students with defective study or test-taking skills and high test anxiety would profit from a combined intervention program to improve their skills and decrease test anxiety.

4.3.4. Basing Treatment on the Broader Diagnostic Picture and Specific Goals of Therapy

The choice of which therapy to use will be influenced not only by the diagnosis of the specific nature of the client's problem and type of test anxiety but also by the broader diagnostic picture, the immediate and long-term goals of treatment, and the therapeutic orientation adopted. For example, although relaxation may not increase the performance of the test-anxious student with study skill deficits, it may be prescribed by the therapist in order to help the student reach the immediate goal of achieving control over test anxiety as a first step toward academic problem solving. Thus, after the anxiety that interferes with learning new study skills is removed, the next step would be training the student in efficient study skills. Furthermore, there are different ways that a therapist may view his or her test-anxious clients' problems (distorted thinking styles, poor problem-solving skills, etc.), and each of these views may give rise to different treatment procedures.

4.3.5. Consideration of Individual Differences

Before implementing a particular treatment, one needs to determine to what extent the treatment may interact with particular client characteristics. For example, some interventions may reduce anxiety or successfully increase the performance of high test-anxious

individuals but have a negative effect on the anxiety performance of others who are low in test anxiety. Also, whereas some people might considerably benefit from relaxation training, experiencing a substantial decrease in anxiety, some have difficulty in acquiring relaxation skills and benefit little from relaxation training. Some clients may even experience relaxation-induced anxiety during relaxation training.

4.3.6. Addressing Multiple Modalities and Loci of Therapeutic Impact

The various components of test anxiety must be dealt with if the anxiety experienced in test situations is to be reduced and improved grades are to be obtained as a result of treatment. It is important that interventions be sufficiently complex to deal with the major facets (cognitive, affective, and behavioral) of the test anxiety experience. Indeed, a treatment would be expected to be most effective if it impacts upon the entire range of components and chain of events leading to anxious manifestations in evaluative situations (arousal, worry, meaning system, internal dialogue, behavioral acts, etc.) rather than focusing on only one aspect of the process.

4.3.7. Interaction among Components of Test Anxiety

A basic consideration is that test anxiety is more than a combination of physiological arousal, negative self-preoccupation, a deficit in stress-related coping skills, and poor study habits. It is the complex interaction among these diverse components that seems to define test anxiety. Because the cognitive, affective, and behavioral components of test anxiety all contribute to the problem of test anxiety and its treatment, it is predicted that an induced change in one system would generally be followed by a change in the other. Thus, therapeutic approaches, which often emphasize cognition, extend to the emotional life and vice versa. For example, it is likely that emotion-focused training (e.g., progressive relaxation) may make the client less anxious and result in a decrease in anxiety-focused, task-irrelevant ideation. Similarly, some forms of cognitive therapy may provide test-anxious subjects with an increased sense of perceived control, which might spill over into the emotional domain and result in lower emotional arousal in a test situation.

It is readily apparent that test anxiety intervention should be based on a careful theoretical analysis of the nature of test anxiety and its key components and manifestations. Traditionally; however, test anxiety treatment studies have mainly evolved from interest in specific behavioral treatment techniques rather than from an analysis of the nature and effects of test anxiety. Indeed, most investigators who have applied behavioral methodology to the reduction of test anxiety have generally given little attention to relating the treatment process to important theoretical conceptions. The current diversity of test anxiety treatments, although supplying the clinician with a rich variety of treatment options to choose from in rendering services, reflects a state of uncertainty marked by the lack of consensus regarding the most effective method for treating test anxiety.

See Also the Following Articles

Academic Interventions ■ Achievement Tests ■ Assessment and Evaluation, Overview ■ Clinical Assessment ■ Coping ■ Educational and Child Assessment ■ Intelligence Assessment ■ Reading Interventions ■ Stress

Further Reading

Ackerman, P. L., & Egested, E. D. (1997). Intelligence, personality and interests: Evidence for overlapping traits. *Psychological Bulletin*, *121*, 219–245.

Covington, M. V. (1992). *Making the grade*. New York: Cambridge University Press.

Eysenck, M. W. (1992). *Anxiety: The cognitive perspective*. Hove, UK: Erlbaum.

Hembree, R. (1988). Correlates, causes, effects, and treatment of test anxiety. *Review of Educational Research*, *58*, 7–77.

Phillips, B. N. (1978). *School stress and anxiety: Theory, research and intervention*. New York: Human Sciences Press.

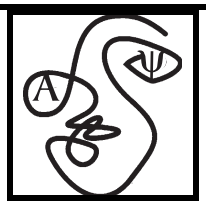
Sarason, S. B., Davidson, K. S., Lighthall, F. F., Waite, R., & Ruebush, B. K. (1960). *Anxiety in elementary school children*. New York: Wiley.

Sieber, J. E., O'Neil, H. F., Jr., & Tobias, S. (1977). *Anxiety, learning and instruction*. Hillsdale, NJ: Erlbaum.

Spielberger, C. D., & Vagg, P. R. (Eds.). (1995). *Test anxiety: Theory, assessment, and treatment*. Washington, DC: Taylor & Francis.

Zeidner, M. (1998). *Test anxiety: The state of the art*. New York: Plenum.

Zeidner, M., & Matthews, G. (In press). Evaluative anxiety: Current theory and research. In A. Elliot, & C. Dweck (Eds.), *Handbook of competence and motivation*. New York: Guilford.



Traffic Safety Assessment

Kåre Rumar
Sturefors, Sweden

1. Introduction
2. Road Safety Problems
3. The Role of Psychology in Road Safety Work
4. Three Ways to Reduce Human Errors in Road Transport
5. Conclusion
Further Reading

GLOSSARY

accident An unavoidable event resulting from circumstances presumed to be beyond control.

active safety An approach to traffic risk management that aims at preventing crashes via improved road and vehicle design, driver education and training, behavioral rules (e.g., speed limits), control and enforcement, etc.

crash A collision between two objects in traffic. By understanding that human control, human decision making, and human decision-making error contribute to circumstances surrounding crashes, we allow for direct involvement (research and education) in the reduction of traffic fatalities and injuries.

passive safety An approach to traffic risk management that aims to minimize the harmful results of crashes via seat belts, air bags, energy-absorbing front ends, crash barriers, etc.

This article describes the scope of the traffic safety problem throughout the world. More than 90% of all transport fatalities in the world occur on roads. Approximately 1 million people are killed on the roads annually, and a majority (85%) of these deaths occur in developing countries. In-depth analyses of crashes show that human error is a major contributing factor in most crashes. Therefore, psychology should be able to contribute to effective countermeasures. The existing road safety problems are analyzed in different ways. The role of psychology is described in association with a general cube model, the role of the driver, and the severity of the road safety problem. Finally, three possible ways to reduce human error in road traffic based on psychological knowledge are described.

1. INTRODUCTION

Worldwide, approximately 1 million people are killed in road traffic each year, and the number increases every year. In comparison with road traffic, the other transport modes (air, sea, and rail) are relatively safe. More than 90% of all transport fatalities occur on the roads. Road safety statistics are very unreliable, except those for fatalities. Therefore, this articles focuses on road traffic crashes and on fatalities. A majority of these fatalities (85%) occur in developing countries. In the developed world, the largest fatality group is occupants of motor vehicles. In the world as a whole, and especially in developing countries, pedestrians comprise the largest fatality group. Contrary to most other causes of death, road traffic fatalities affect mainly young persons. Therefore, road traffic fatalities result in the highest number of lost years of all causes of death in the developed world. The World Health

©2004 Elsevier Inc.
All rights reserved.

Organization estimates that road crashes are the third leading health threat considering lost years and lost working time due to injuries. Thus, road crashes are a major public health problem throughout the world.

In most countries, large crashes that occur in the air, on the sea, or on rail are studied in-depth to determine the causes and to prevent similar crashes from occurring again. These studies have been quite successful and have resulted in the prevention of many crashes and fatalities. Efforts to study the causes of all fatalities on the roads by means of in-depth accident investigations have not been as successful. This process is, for many reasons, much more difficult. For example, it is much more difficult than it is with air, sea, or rail crashes to establish the important events that occurred immediately before the crash. However, the results of investigations of road crashes show convincingly that human error is a major contributing factor in more than two-thirds of the crashes (Fig. 1). Considering that human error plays such a major role in road crashes, it is evident that psychology is a discipline that may have much to contribute to road safety work.

In this article, the term crash is used instead of the traditional term accident because accident implies that the event is totally random and cannot be avoided. Consider the statements, "It was a complete accident" or "It happened by accident." By labeling road traffic crashes as accidents, it is psychologically more difficult to convince people to contribute to the avoidance of crashes.

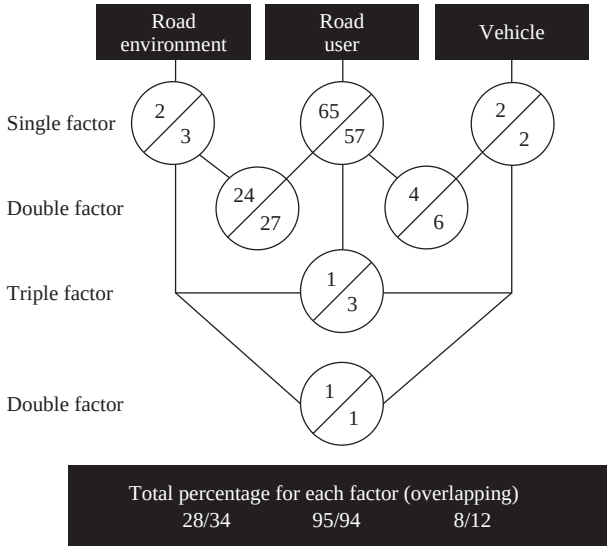


FIGURE 1 Percentages of the main factors contributing to road crashes according to the two largest crash studies (United States and the United Kingdom).

2. ROAD SAFETY PROBLEMS

There are many ways to describe road safety problems. Here, two examples are presented.

2.1. Exposure–Crash Risk–Injury Consequence

A fruitful way to describe road safety problems is to use what is sometimes called the Swedish ice cube (Fig. 2). The number of people killed in an area depends mainly on three variables:

- Exposure in traffic (number of kilometers, hours, trips, etc.)
- Probability of a crash given a certain exposure (crash risk or rate)
- Probability of being killed given the crash (crash consequence)

These three variables multiply in a complex way to produce the total number of people killed. By multiplying the variables, a volume (cube) is obtained. The size of this volume describes the severity of the road safety problem (how many people are killed). Consequently, when one wishes to reduce this volume, the three variables are employed. By reducing one, two, or three of these variables, the number of fatalities will be reduced.

2.2. Primary, Secondary, and Tertiary Problems

Another way to describe the road safety problem is to use the various specific road safety problems as starting points. The most well-known problems are alcohol, speeding, young drivers, low usage of protective devices (e.g., seat belts and helmets), tailgating, darkness, low friction, etc. If one can find the appropriate countermeasure for each one of these problems, one may think that the overall road safety problem has been solved. Unfortunately, this is not the case. There are solutions for many of these primary problems. However, for many measures, the effect is much less than expected. Another problem is that the known countermeasures are not implemented. These primary problems (alcohol, speed, etc.) are therefore only

components of the overall problem. Thus, one has to look elsewhere. One has to search for the obstacles that reduce the effect of any countermeasure or that prevent known countermeasures from being implemented. Consequently, road safety problems can be described on three levels as follows:

- Primary problems, which require countermeasures (alcohol, speeding, young drivers, etc.)
- Secondary problems, which reduce the effect of any countermeasure (inefficient enforcement, inadequate driver training and licensing, inadequate information, etc.)
- Tertiary problems, which prevent countermeasures from being implemented (low value of road safety measures, no defined goals, no quantitative targets, etc.)

3. THE ROLE OF PSYCHOLOGY IN ROAD SAFETY WORK

In this section, the two methods of defining road safety problems are used to analyze the role that psychology may play in the road safety process from problem analysis via research to implementation and follow-up.

3.1. The Cube Model

The ice cube model (Fig. 2) clearly shows that traffic psychology has much to contribute to all three variable of the model. Along the exposure axis, the task could

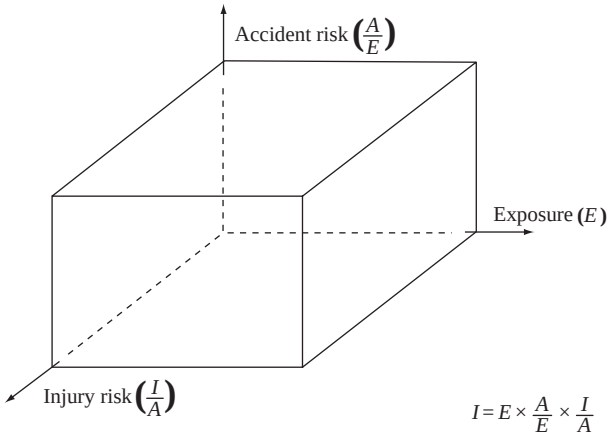


FIGURE 2 The magnitude of safety problems (number of human injuries and fatalities; *I*) as a function of the product of three variables: exposure (*E*), crash risk (*A*/*E*), and injury consequence (*I*/*A*).

be, for instance, to determine what makes people choose one transport mode over another. It is known that public transport is at least 10 times safer than private vehicles. What is the relative influence of price, time, comfort, freedom, company, etc. on modal choice? Results indicate that price and safety level are relatively unimportant, whereas freedom and comfort are very important.

Regarding the crash risk axis, psychology has always been one of the main contributors to understanding. This axis could also be called crash prevention or active safety, both of which could be improved by reducing the number of human errors. Traditional approaches to crash prevention are to improve roads, vehicles, and the driver or other road users.

In traditional road safety work, three E's are often mentioned, referring to engineering, education, and enforcement. The first E has already been mentioned (roads and vehicles). The other two E's are of a more psychological nature. Education may include general information, directed campaigns, education in schools, or special courses—all for the purpose of training various types of road users, such as pedestrians, bicyclists, motorcyclists, and drivers of passenger cars or heavy vehicles. Also, enforcement has many forms, such as hidden or open, intense or rare (probability), reliable or random, immediate or delayed, and manual or automatic.

Education and enforcement methods have been used since the introduction of automobiles. However, information, campaigns, education, and training have not provided the expected positive road safety results. Enforcement, on the other hand, has proved very effective. The problem is that it is very costly if carried out manually. This is why modern technology (e.g., automatic cameras and electronic identification of cars) is being used more frequently.

In recent decades, increasingly more resources have been used in human factor efforts. This means that the technical environment (roads, vehicles, and traffic legislation and control) should be adapted to human characteristics. This approach is discussed later.

Work along the third model axis tries to reduce crash consequences. This approach is often called passive safety. This approach has been very successful in recent decades. Seat belts, air bags, energy-absorbing front ends, and crash barriers are good examples of this approach. Other efforts include improved rescue service and treatment of injured persons. Along this axis, psychology has not contributed as much. However, crash protective devices (seat belts and helmets) have no effect if they are not used. The following are typical

questions for psychologists: Why do different groups of road users not use available crash protective devices more often than they do? What should be done to increase the usage rate of crash protective devices?

3.1.1. Specific Psychological Problems

Work along the second model axis (active safety) has not been as effective as originally hoped. One of the reasons is that the reactions of drivers under changing conditions are difficult to predict. For instance,

- How will drivers use better visibility, better friction, and improved brakes? Will they drive more safely, increase their speed, or both?
- What is the driver really doing when he or she is driving? What should a useful driver model look like?

The first problem deals with what has often been called the risk compensation hypothesis. This hypothesis states that drivers tend to stay at a constant risk level irrespective of various changes in circumstances. In order to change the risk level, the basic values of the driver have to be influenced. Experience shows that the more feedback on safety improvements the driver has, the more he or she tends to speed, and that the less feedback the driver has, the more he or she tends to drive safely. Street lighting and studded tires on ice are good examples of the first and daytime running lights are a good example of the second situation.

The second problem (designing a useful driver model) has occupied transportation psychologists for at least 70 years. It is fair to say that there is no generally useful driver model. However, there are many driver models that work quite well for limited tasks (e.g., tailgating).

3.2. Three Problem Levels

The second way to describe the road safety problem is to divide it into three levels (primary, secondary, and tertiary). Traffic psychology is one of the major disciplines working to find effective countermeasures to the primary problems (alcohol, speeding, young drivers, etc.). The success rate has not always been very good. Two of the major reasons for the reduced levels of effect are that enforcement levels have been inadequate and the awareness of the seriousness of the road safety problem has been too low. Driver motivation to follow recommendations and legislation has therefore been low.

Regarding the second- and third-level problems, traffic psychology has been practically the only contributor to solutions. By refining education and training and by improving enforcement levels and strategies, traffic psychology has contributed to increasing the efficiency of most countermeasures. In both cases, intelligent transportation systems have proven to be a potentially very effective tool.

The third-level problems are of a very psychological character. The idea is mainly to increase the motivation for increased road safety. Politicians should be motivated to implement a number of road safety measures, citizens should be motivated to ask for the implementation of road safety measures, and drivers should be motivated to conform to the safety recommendations and legislation. Sweden was the first country to try to work systematically on this problem level. The Swedish "Vision Zero" has proven very effective in reducing the motivational problems, especially on the political level. Vision Zero states that nobody following the rules should be killed or impaired in road traffic. Responsibility for the design of the system lies with the road authorities, who own the system.

Third-level problems also include the way road safety work is monitored and followed up. Without feedback, mistakes may be, and have been, repeated. General information campaigns are one of the most popular road safety measures among politicians and other decision makers. The few follow-up studies that have been carried out show that the road safety effects of such general campaigns are, at best, very limited.

By establishing quantitative targets for the various sectors of road safety work, results can be followed up easily by a kind of result management. The targets must not be set only in terms of crashes and fatalities because these are unreliable indicators of road safety in the short term. In order to obtain a quick follow-up system, performance indicators are used in modern road safety work. Such performance indicators are often behavioral by nature, such as the speed distribution of cars, the time gaps between cars, and the proportion of drivers using seat belts. Here, psychology can contribute substantially.

4. THREE WAYS TO REDUCE HUMAN ERRORS IN ROAD TRANSPORT

Assume that in a large majority of road crashes, human error is a major contributing factor (Fig. 1). One of the priority targets in road safety work should be to reduce human errors. To do so, the following three

approaches must be considered, to which psychology can make potential major contributions:

- Improve selection of drivers
- Modify (improve) driver behavior
- Adapt the traffic conditions and driver tasks to human characteristics

The first approach (Fig. 3) is to select drivers in such a way that those who do not perform well enough and who often commit errors will not be allowed on the roads. Currently, drivers are selected on the basis of medical, knowledge, and skill tests. These tests and their passing criteria are different in different countries. It is difficult to make these tests more stringent and tough because driving a car is considered almost a human right in most countries. The situation for pilots, for instance, is very different. Selection procedures and criteria are to a very large extent of a psychological nature (e.g., test validity and reliability).

The second approach (Fig. 3) is to try to improve the knowledge, performance, attitudes, and behavior of drivers. Typical methods include training, education, information, and enforcement. In order to carry out this improvement, considerable knowledge concerning learning psychology is required. Also, it is important to avoid the impairment of driver behavior due to fatigue, alcohol, and drugs.

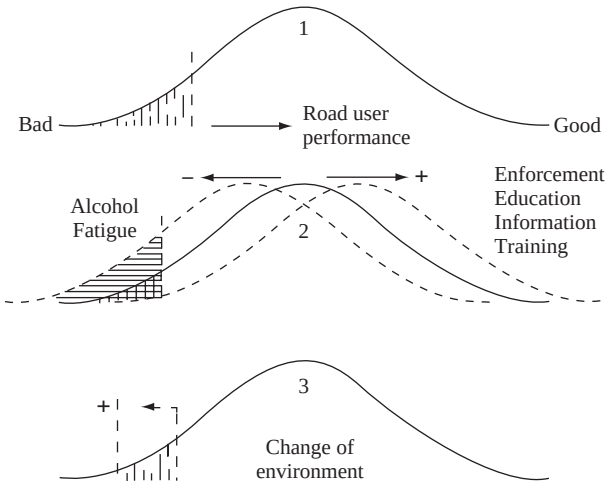


FIGURE 3 Three ways to reduce human error: (1) restrict road users who make many errors; (2) improve the performance of road users through information, education, training, legislation, and enforcement while ensuring that their performance is not impaired by negative factors, such as alcohol, drugs, and fatigue; and (3) adapt the environment to human characteristics so that it becomes easier to walk, bicycle, and drive.

The ancient Greeks stated that "Man is the measure of everything." In other words, if human beings commit errors in road traffic, there is something wrong with the road traffic system. The third approach (Fig. 3) means that the traffic system should be adapted to human characteristics and limitations, mainly concerning human perceptual and cognitive characteristics (the ability to perceive and interpret the traffic scene as a road user). The traffic systems—the road, road environment, road lighting, traffic control devices, vehicle dynamic characteristics (stability, suspension, braking, steering, etc.), and vehicle exterior and interior systems and details (lighting, rearview mirrors, controls, instruments etc.)—should be designed in such a way that they are easy to decode, interpret, and understand. In order to carry out such an adaptation, qualified psychological knowledge and research are required.

5. CONCLUSION

The previous discussion shows how both ways of describing the road safety problem lead to counteractions, which to a large extent are comparable. Therefore, it is not crucial which model is used. However it is crucial that the involved actors can understand and share the model. A common road safety model simplifies cooperation between all road safety actors. This article also notes the major role that human error plays in road crashes. This is a strong indication of the importance of traffic psychology in past and present road safety work, and it indicates that traffic psychology will be even more important in future road safety work.

See Also the Following Articles

Accidents in Transportation ■ Aviation ■ Driving Safety ■ Models for Transportation ■ Transportation Systems, Overview ■ Travel Behavior and the Environment

Further Reading

American Automobile Manufacturers Association. (1998). *World motor vehicle data*. Washington, DC: Author.

Barjonet, P.-E. (Ed.). (2001). *Traffic psychology today*. Boston: Kluwer.

Brown, I., & Groeger, J. (Eds.). (1990). Errors in the operation of transport systems. *Ergonomics*, *33*(10/11).

Evans, L. (1991). *Traffic safety and the driver*. New York: Van Nostrand Reinhold.

Evans, L., & Schwing, R. C. (Eds.). (1985). *Human behavior and traffic safety*. New York: Plenum.

International Road Federation. (2000). *World road statistics*. Geneva: Author.

Näätänen, R., & Summala, H. (1976). *Road user behaviour and traffic accidents*. Amsterdam: North Holland.

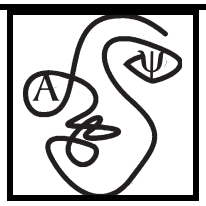
Rothengatter, J. A., & de Bruin, R. A. (Eds.). (1987). *Road users and traffic safety*. Maastricht, The Netherlands: Van Gorcum.

Rumar, K. (1999). *The first European transport safety lecture*. Brussels: European Transport Safety Council.

United Nations. (2000). *Traffic safety: Problems and countermeasures*. Geneva: United Nations.

World Bank. (2001). *Why are road crashes a problem?* Geneva: Global Road Safety Partnership.

World Health Organization. (1999). *Injury—A leading cause of the global burden of disease*. Geneva: World Health Organization.



Training, Cross-Cultural

David J. Dalsky
University of Mississippi, Mississippi, USA

Dan Landis
University of Hawaii, Hilo, Hawaii, USA

1. The Need for Cross-Cultural Training
2. People Involved in Training Programs
3. Structure of Training Programs
4. Program Evaluation and Dimensions of Success
5. Conclusion
   Further Reading

GLOSSARY

cultural assimilator A series of critical incidents involving cultural misunderstandings that are typically used in cross-cultural training programs.

didactic technique A training method that includes lecture or supplying information through other means in which trainees take a passive role in learning.

diversity training A type of cross-cultural training whose goal is to reduce prejudice and discrimination in settings within a country.

experiential technique A training technique that involves a cultural simulation, role play, or related activity in which participants play an active role.

host country The country in which the trainee is preparing to live for a period of time.

intercultural sensitivity The degree of a sojourner's ability to recognize cultural differences in norms and values.

predeparture training Sessions of a training program that occur before the sojourner leaves for his or her host country and/or immediately upon arrival to the host country.

reentry training Sessions of a training program that take place just prior to the departure of the sojourner from his or her host country and/or upon his or her return home.

sojourner One who takes a temporary residence abroad.

Cross-cultural training involves preparing sojourners to be successful in their host country. Trainers commonly have educational backgrounds in the social or applied sciences (e.g., psychology, communications, and counseling) or a related field, and trainees are typically undergraduate study-abroad students, international business employees, diplomats, or volunteers with international assignments. Although the content of each cross-cultural training program varies widely, the essential goals are to introduce cultural differences and adaptive strategies for cultural adjustment through the use of a variety of training techniques, including lecture, discussion, and experiential learning activities that involve the active participation of trainees. A successful and effective training program will include a good balance of these techniques and will allow trainees to reflect upon their own experience in the program's activities through discussion and cultural theory integration from lecture. Depending on the participants' characteristics, training programs may be specifically designed for the host country of the trainees or generally geared for adapting to life abroad. More complete programs involve predeparture and reentry sessions, and trainers maintain contact with participants during their stay abroad. Upon completion, cross-cultural training programs are evaluated and continuously improved for the benefit of future trainees. Although the majority of programs involve sojourners, cross-cultural training programs can also be structured to help people communicate and work more effectively with members of different cultures, races, or ethnic groups living in their own country.

©2004 Elsevier Inc.
All rights reserved.

1. THE NEED FOR CROSS-CULTURAL TRAINING

Cross-cultural training programs are typically designed to introduce sojourners to cultural differences and to prepare them for success in their host countries. Successful programs are tailored to the needs and goals of the trainees, which are assessed prior to the start of the training program. Especially when based on an underlying theory, well-designed training programs have been shown to be more effective than no training at all or the brief, less structured country-specific orientation sessions typically held by international companies, volunteer organizations, or university study abroad programs.

When the purpose of the training program focuses on improving within-country interactions, the emphasis is placed on reducing stereotypes and prejudices. This type of cross-cultural training, also known as diversity training, is becoming increasingly more common in the United States as the population continues to become more diverse and work groups consist of employees from various racial, ethnic, and cultural backgrounds.

1.1. Needs Assessment

The needs assessment typically involves surveying trainees for information regarding their family, work, language, or other characteristics. Trainees with families will be concerned about locating a suitable living place, schools for their children, and a place of employment for their spouses. Details involving their place of employment must also be addressed, including the use of a foreign language in the workplace or at home with a host family, change of position in a company, transportation to and from work or school, and other specific concerns. Trainees will often be faced with quite different living experiences in their host country than those to which they are accustomed. Particularly not to be overlooked is any reduction in status that the overseas position may involve, making it even more difficult for the sojourner to adjust to his or her new role and life in the host country. Adaptive strategies to prepare for and handle these changes are introduced and practiced in cross-cultural training programs.

Diversity training program trainers consider the needs of trainees and the organization to which they belong. Work groups must be surveyed for demographic information and the trainer must be familiar with the demands and goals of the particular groups. Organizational leaders also need to be consulted for the identification of any issues that may be of special concern for the trainer to keep in mind when structuring his or her program.

1.2. Goal Identification

The identification of goals is essential for the success and effectiveness of any cross-cultural training program. Specific goals are based on the particular needs of the trainees involved and should be expressed to them by the trainer at the start of the program. If the goals of the program are not aligned with needs of the trainees, participants can be uncooperative or doubtful of the usefulness of the program. Statistics on the percentage of sojourners successfully completing their overseas assignments can also be presented in support of the benefits of cross-cultural training programs. Once participants are convinced of the need for training, they are usually more willing to play an active role in the program. An effective cross-cultural training program usually combines learning exercises and discussions with lectures on culture theory. Some of the most general goals of cross-cultural training programs include increasing awareness of cultural differences, identifying and improving attitudes, and introducing new behaviors to aid in adaptation to the host culture.

1.2.1. Increasing Awareness of Cultural Differences

One major goal of cross-cultural training programs involves the trainees gaining an increasing awareness for the norms, values, and other specific information about their own culture and the host culture. Its importance can be demonstrated through specific examples of cross-cultural misunderstandings resulting, in part, from a lack of awareness of cultural differences. This goal of the program can be achieved through the presentation of information in lecture or video format and reinforced through activities and discussion.

1.2.2. Identifying and Improving Attitudes

Helping the participants of a cross-cultural training program to become aware of preexisting negative attitudes or biases toward interacting with culturally different individuals can be difficult but is a very important goal. It is difficult because people do not generally like to be made aware of their own negative self-aspects, but it is important for these negative

attitudes to be exposed and identified before sojourners must interact with members of their host culture.

A cultural simulation activity is one of the best techniques to use for the participants to get a feel for what it is like to face intercultural communication problems. Through these activities, they will be immersed in a situation in which they will most likely experience anxiety, frustration, and anger. After these are identified in a discussion, coping techniques can be suggested and practiced.

1.2.3. Introducing New Behaviors

Particularly when the host culture is very different from that of the trainees, one of the goals of the program may be to introduce and practice new behaviors that are appropriate and very common in the host culture. Exercises involving greetings may be especially useful for a Westerner who is planning a sojourn to an Eastern culture, for example. These exercises may be practiced in a role play preferably with a member of the host culture who is involved in the training program.

2. PEOPLE INVOLVED IN TRAINING PROGRAMS

A cross-cultural trainer plays many roles during the course of a training program. He or she must be very sensitive to the needs of the trainees, sometimes acting as a teacher, counselor, and role model. Trainees must also be open to the experiences of the program and willing to express their feelings and voice their opinions. Other members of a cross-cultural training program include members of the host culture or returning sojourners who may act as role models and cofacilitators during the program.

2.1. Trainers

The leaders of cross-cultural training programs typically have educational backgrounds in the social or applied sciences (e.g., psychology, sociology, anthropology, linguistics, communications, and counseling) along with cross-cultural experiences of their own. These qualifications, however, do not necessarily make for an effective trainer. A trainer with reputable credentials and cross-cultural experience, for example, may lead a program based solely on his or her intuitive feelings of what he or she thinks needs to be done. Since this type of approach has been shown to be less effective than those grounded in psychological or communication theories, it is essential for an effective trainer to be not only well educated and experienced but also familiar with and willing to use relevant theory and research in applicable fields.

Important personal characteristics of effective trainers include empathy, patience, flexibility, open-mindedness, and a high level of intercultural sensitivity. An empathetic and sensitive trainer, for example, is quick to recognize highly ethnocentric trainees and can actively restructure his or her program to suit their needs. Responding to the specific needs of the trainees requires a great deal of patience and flexibility on the part of the trainer, especially when faced with a group that has not voluntarily chosen to sojourn or participate in the training program. After the program has been evaluated, an open-minded trainer will thoughtfully consider the feedback of his or her trainees and make appropriate adjustments for future programs.

2.2. Trainees

Trainees of cross-cultural training programs are typically international business employees, diplomats, volunteers, or university students preparing to study abroad. Since the structure of a program primarily depends on the type of participants involved, business employees, for example, will go though a quite different training program than undergraduates planning to study abroad. Training techniques used for business employees will focus on the particular issues they will be facing in the workplace and with their families, whereas programs for students studying abroad will discuss issues related to host families and schools.

Individual characteristics of the trainees, including degree of ethnocentrism and intercultural sensitivity, also play a role in the structure of the program. Highly ethnocentric individuals may benefit more from experiential activities, whereas those who are less ethnocentric and have more cross-cultural experience may gain more from discussing cultural or communication theories presented in lecture. Programs with a large number of trainees may be structured so the latter individuals serve as group discussion leaders, reporting back to the entire group while the trainer provides feedback.

2.3. Others

Additional participants of cross-cultural training programs may include members of the host culture who

might serve on a sort of discussion panel to answer trainees' questions or actively participate in role plays. These role plays are designed to introduce new behaviors to the trainees, providing an opportunity to "practice" with a member of the host culture and experience a cross-cultural interaction prior to their sojourn. Also, returning members of the same organization who have spent time in the same host culture may be involved in programs to answer trainees' questions, provide advice, and talk about their experiences.

3. STRUCTURE OF TRAINING PROGRAMS

After the goals of the program have been established, trainers must consider the type of training techniques to be used (lecture, discussion, or experiential activity), the focus of these techniques (culture-general or culture-specific), sequence, and time allocation for each. Time and resource availability can pose a difficult challenge for trainers while structuring a cross-cultural training program, and all these elements must be considered when designing a program.

One major structural issue to consider is the inclusion of a reentry component to the training program. The most complete cross-cultural training programs involve a series of predeparture training sessions before the sojourners leave for their host country, sessions immediately upon arrival to their host country, and a series of reentry sessions beginning before the sojourners return home. Reentering is often more stressful than entering a host country for the first time, which is compounded by the fact that sojourners are generally not expecting it to be a difficult transition for them. Reentry training offers a time to discuss expectations with sojourners before they leave their host country and provides a social support network of people that have been through a similar experience. Additional benefits of reentry workshops include building greater insight into self-identity and a time for reflection on what was learned during the overseas experience.

3.1. Types of Training Techniques

The techniques available for use in a training program can be classified into two broad categories. Didactic techniques involve the passive role of trainees and include lectures and videotapes, whereas experiential techniques involve the active role of trainees and include role plays, discussions, cultural assimilators, and cultural simulation activities.

3.1.1. Didactic Techniques

Lectures can introduce essential aspects of psychological or communication cultural theories. A psychological theory that is particularly useful for sojourners from the West who are going to live in the East and visa versa is individualism/collectivism. The defining features of these subjective cultural dimensions can be introduced and examples of common cross-cultural misunderstandings involving them can be given to aid in the understanding of these concepts.

For example, an indirect communication style is one tendency of a collectivistic country's culture, whereas a direct style of communication is more typical in an individualistic country's culture. After a specific example is provided, possibly through the use of a cultural assimilator, trainers can then turn the focus to the underlying theory and offer it to the trainees as a kind of framework for sorting through and understanding their own cross-cultural experiences to be encountered in their host country.

Another useful theory that can be introduced in lecture is the anxiety—uncertainty management (AUM) theory. The essential components of this communication theory revolve around the core components of mindfulness, anxiety, and uncertainty. First, trainees can be taught what it means to be mindful and may be given examples to illustrate its importance in cross-cultural interactions. Next, the trainer might explain that managing anxiety and uncertainty is necessary for successful cross-cultural interactions and demonstrate this through the use of specific examples. Trainees would then be asked to generate some ideas regarding what techniques they can use to manage their own anxiety in their host country. Following this, the trainer can ask them to provide specific examples of some difficult situations they may expect to encounter along with their own strategies for managing anxiety and reducing uncertainty. If this lecture follows a culture simulation activity, trainees can be asked to reflect upon how they felt during the simulation using the AUM theory to frame their responses. The trainees can then practice using these AUM-based strategies in a cross-cultural simulation activity.

3.1.2. Experiential Techniques

In cultural simulation activities, participants are asked to form groups that receive separate instructions

regarding group communication or behavioral rules. The trainees are then asked to interact with members of the other groups for awhile to accomplish some type of task until they experience emotions that typically arise in actual cross-cultural encounters. Examples of such activities most commonly used are BaFa BaFa, Barnga, and Ecotonos. One important source of misunderstandings is that members of the host culture make attributions concerning their own behavior that are very different from the attributions that the sojourners make concerning the behavior of the hosts. Exposing these discrepancies through simulation activities has proven to be beneficial. After the simulation, a trainer may debrief the trainees on the purpose of the activity and follow up with a discussion based on the principles taught in the lectures of the training program.

Another experiential training technique involves the use of cultural assimilators, through which the trainees are asked to work and discuss. These sets of incidents involve common cross-cultural misunderstandings followed by four possible options for the most correct explanations. Those that are grounded in cultural theories can be integrated with information presented in lecture and combined with the participants' experience in simulation activities.

For example, an incident from one such assimilator involves a new colleague arriving from China who occupies a desk next to an American colleague. As time passes, the American begins to be bothered by the strong body odor of his new colleague and wishes to communicate this with him. The assimilator then presents a variety of possible responses. Among these are more or less direct approaches of how the American should communicate with the Chinese colleague. Culturally insensitive Western trainees are likely to choose a more direct response, such as "Present deodorant to the colleague and ask him politely to use it," rather than a more indirect approach, such as "Talk to the colleague's closer friends or family and have them communicate the message to him." This assimilator based on individualism/collectivism theory offers explanations based on the frequency of direct and indirect responses in the two cultures, thus identifying which explanation is most likely to be correct. In this example, the trainee is exposed to the importance of "saving face" in more collectivistic countries and the necessary indirect communication style that is characteristic of this value. Like most assimilators, this particular incident is available in a written format, but it is also available in an interactive computer audio/video format. Paper has traditionally been the only assimilator format available; however, computer-based software programs have recently become available.

The role play is another example of an experiential training technique. Particularly when individuals from the host country are involved in a training session, a role play can be an excellent opportunity for the trainees to practice behaviors that may not be practiced as often or at all in their own culture. An American sojourner whose host country is Japan, for example, may be placed in a situation with a Japanese playing the part of a business colleague. They might be asked to play out a situation in which they are meeting for the first time, and the trainee would then practice exchanging business cards, bowing, and using formal Japanese when asking a "new Japanese business colleague" appropriate questions.

Within experiential and didactic training techniques, another distinction differentiates between those focusing on a specific host culture, called culture-specific, and training that focuses on broad aspects of cultural differences, called culture-general. Among the advantages of the cultural-general approach is the ease of accessibility to a set of culture-general assimilators that have been published or other sets that are in wide circulation. Culture-specific assimilators, although very useful for groups of trainees sojourning to the same country, can be difficult to obtain without having a large network of professionals in the field.

3.2. Time and Resource Availability

Although trainees would benefit most from training programs wide in breadth and depth, time constraints and limited resources often pose a challenge for trainers to make the most of what is available to them. An ideal program would include thorough, informative lectures combined with language training (if applicable), cultural simulation activities, and discussions to integrate all the experiences in the training program. The focus (culture-general or -specific) would be designed to suit the needs of the individuals in the group and, at a minimum, 3 full days of training is generally recommend for a training program to be effective.

4. PROGRAM EVALUATION AND DIMENSIONS OF SUCCESS

Anonymous written feedback from participants is perhaps the most basic means of evaluating how well the program was received by the trainees. This feedback

may be used to modify future programs and is an indication of how well the trainer performed in his or her role as the facilitator of the group. Regarding particular training techniques, the results of laboratory studies support the effectiveness of assimilators with regard to raising intercultural sensitivity and reducing prejudices. Results of longitudinal studies also support the effectiveness of cultural assimilator-based training on cross-cultural adjustment, communication, and interpersonal problem-solving ability. These three dimensions of success in the host culture are generally known as personal adjustment, interpersonal effectiveness, and professional effectiveness.

4.1. Personal Adjustment

During the first month or two of the sojourn, individuals typically endure the uncomfortable experience known as culture shock. However, after this initial negative reaction to the new environment and people, those who have completed cross-cultural training programs have been found to experience less stress in their host culture regarding their personal adjustment in comparison to a group of sojourners that did not receive formal cross-cultural training. Information about cultural differences that is introduced in the training program combined with their simulated cross-cultural experiences and rehearsed behaviors allow them to anticipate difficulties and adapt more readily to situations in their host culture.

4.2. Interpersonal and Professional Effectiveness

The two remaining dimensions of success in the host culture are more closely related to the sojourners' interactions with their new friends and colleagues. These are largely a function of the sojourner's ability to communicate in a second language and his or her familiarity with the host culture's norms and values. In the training program, the use of cultural assimilators, role plays, and the presentation of information related to the specific work environment can prepare a sojourner to do well on this dimension. Indeed, compared to a group with no formal training, trainees are more likely to complete their overseas work assignments and have a better social support network in the host culture.

5. CONCLUSION

Well-designed cross-cultural training programs can better prepare sojourners for their experience in their host country and experience upon returning to their native country. A successful program lies in the alignment of goals to the needs of the participants and the rapport created between the trainer and the trainee. An effective program depends primarily on the blend of didactic and experiential techniques involving the use of culture-general or culture-specific training or a combination of the two. It is the job of the trainer to blend the optimal mix of the two types in order to satisfy the participants of the program while maximizing its effectiveness.

See Also the Following Articles

Diverse Cultures, Dealing with Children and Families from ■ Interpersonal Behavior and Culture ■ Social Skills Training

Further Reading

Brislin, R., & Yoshida, T. (1994). *Intercultural communication training: An introduction*. Thousand Oaks, CA: Sage.

Cushner, K., & Brislin, R. (1996). *Intercultural interactions: A practical guide* (2nd ed.). Thousand Oaks, CA: Sage.

Gudykunst, W. B., & Kim, Y. Y. (2003). *Communicating with strangers* (4th ed.). Boston: McGraw-Hill.

Kim, Y. Y., & Gudykunst, W. B. (Eds.). (1988). *Cross-cultural adaptation: Current approaches*. Newbury Park, CA: Sage.

Landis, D., Bennett, M., & Bennett, J. (Eds.) (2003). *Handbook of intercultural training* (3rd ed.). Thousand Oaks, CA: Sage.

Landis, D., & Bhagat, R. (Eds.). (1996). *Handbook of intercultural training* (2nd ed.). Thousand Oaks, CA: Sage.

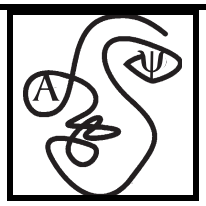
Landis, D., & Brislin, R. (Eds.). (1983). *Handbook of intercultural training* (Vols. 1–3). Elmsford, NY: Pergamon.

Nipporica Associates. (1993). *Ecotonos*. Yarmouth, ME: Intercultural Press.

Shirts, R. G. (1974). *BaFa BaFa: A cross-cultural simulation*. Del Mar, CA: Simulation Training Systems (formerly Simile II).

Thiagarajan, S., & Steinwachs, B. (1990). *Barnga*. Yarmouth, ME: Intercultural Press.

Triandis, H. C. (1995). *Individualism and collectivism*. Boulder, CO: Westview.



Traits

Jesús F. Salgado
University of Santiago de Compostela, Santiago de Compostela, Spain

1. Traits as Human Characteristics of Individual Differences
2. The Big Five Dimensions of Personality
3. Criticism of the Trait Approaches and the Five-Factor Model of Personality
4. Traits, Heritability, and Genetics
5. The Big Five and Organizational Behavior
Further Reading

GLOSSARY

agreeableness The degree to which individuals are cooperative, warm, and agreeable versus cold, disagreeable, rude, and antagonistic.

conscientiousness The degree to which individuals are hard-working, organized, dependable, reliable, and persevering versus lazy, unorganized, and unreliable.

emotional stability The degree to which individuals are secure, nonanxious, calm, self-confident, and cool versus insecure, anxious, depressed, and emotional.

extraversion The extent to which individuals are gregarious, assertive, and sociable versus reserved, timid, and quiet.

non-shared environment Environmental factors that do not contribute to sibling similarity.

openness to experience Defines individuals who are creative, curious, and cultured versus practical with narrow interests.

shared environment Environmental factors that contribute to twin and sibling similarity.

traits Dimensions of individual differences in tendencies to show consistent patterns of thoughts, feelings, and behaviors.

This article describes the concept of traits as they are used for describing the human personality and the relationship between the traits and human languages. It also discusses the psycholexical approaches for describing and explaining personality. The most accepted trait-based personality model, the five-factor model, is described and some criticisms are presented. Additionally, recent evidence of heritability and genetic bases of the big five personality dimensions is discussed. Finally, advances regarding the relationship between the big five personality dimensions and organizational behaviors and processes are mentioned and the magnitude of the relations is reported.

1. TRAITS AS HUMAN CHARACTERISTICS OF INDIVIDUAL DIFFERENCES

It is very probable that human beings have used personality traits for describing themselves and others since they were first able to verbally communicate. This statement suggests four relevant implications: (i) Traits are conceptual attributes used for describing ourselves and others, (ii) traits are only possible because they are rooted in human language, (iii) traits are stable over time and across situations, and (iv) traits are related to human behavior.

The trait-based approaches to personality focus on the conceptualization that individuals differ from one another in a small number of dimensions that are

©2004 Elsevier Inc.
All rights reserved.

stable over time and across situations. In this sense, traits are defined as dimensions of individual differences in tendencies to show consistent patterns of thoughts, feelings, and behaviors that allow the characterization of human beings. Tendencies implies that traits are dispositions and not absolute determinants of human action. Consistent patterns of thoughts, feelings, and actions implies that traits should be distinguished from habits, repetitive actions, and mechanical behaviors. Furthermore, as Allport noted many years ago, personality traits are not real but rather abstractions that cannot be directly measured and must be inferred.

Different approaches to personality traits have been suggested throughout the years; the first was the lexical approach to the traits suggested by Galton in the 19th century. Galton proposed that personality characteristics are encoded in the human language and, consequently, are included in dictionaries. Following Galton's seminal work, in 1936 Allport and Odbert classified all the trait words included in the American English language as were registered in *Webster's New International Dictionary*. They identified 17,953 personality terms classified in four categories: personal traits, evaluative traits, methaphorical and temporary states, and moods. After eliminating all the words included in the last three categories, they developed a list of 4504 words that Americans used to describe themselves and others. Because this large list was not very practical from both a theoretical and an applied standpoint, in 1943 Cattell tried to reduce the number of trait words using a variety of conceptual and statistical techniques. First, he reduced the large number of words to 171 using conceptual and semantic similarity. Then, using cluster analysis, he grouped the short list of words and found 35 clusters. A factor analysis of the 35 bipolar scales produced 12 primary dimensions. Cattell added four dimensions that he had found by analyzing questionnaires and the result was the basis of his theory of personality: 16 primary personality factors. Throughout the years, a number of researchers have investigated Cattell's structure of personality and found some incongruence.

Soon after Cattell's study, Eysenck stated his theory of personality, which consisted of two independent factors, namely neuroticism (the negative pole of emotional stability) and extraversion. Later, Eysenck added a third factor, psychoticism, and the result was the well-known P-E-N system.

2. THE BIG FIVE DIMENSIONS OF PERSONALITY

The psycholexical approach to personality has led to the accumulation of empirical evidence supporting five personality dimensions that summarize the relations between the personality traits. This evidence was found for many different nations and with many different languages, although some minor differences exist across the national borders and the languages with different linguistic roots. These five personality dimensions, the "Big Five" factors, are extraversion, agreeableness, conscientiousness, emotional stability, and intellect. These five factors have also been labeled surgency (extraversion), neuroticism (emotional stability), openness to experience (intellect), dependability (conscientiousness), and many other terms. Currently, the most used labels are emotional stability, extraversion, openness to experience, agreeableness, and conscientiousness. The five-factor model (FFM) assumes that the personality traits are hierarchically arranged, with specific narrow traits at lower levels in the hierarchy and the broad trait dimensions at the top.

The current FFM of personality originated from works by Fiske, Borgatta, Norman, Smith, and Tuppes and Christal, who factor analytically reproduced a highly stable structure with five factors. However, they were unable to reproduce Cattell's 16 primary factors, and they found evidence for only five factors. In the FFM, emotional stability concerns the degree to which the individual is secure, nonanxious, calm, self-confident, and cool versus insecure, anxious, depressed, and emotional. Extraversion concerns the extent to which individuals are gregarious, assertive, and sociable versus reserved, timid, and quiet. Openness to experience defines individuals who are creative, curious, and cultured versus practical with narrow interests. Agreeableness concerns the degree to which individuals are cooperative, warm, and agreeable versus cold, disagreeable, rude, and antagonistic. Conscientiousness concerns the degree to which individuals are hardworking, organized, dependable, reliable, and persevering versus lazy, unorganized, and unreliable.

The five factors represent the common variance among a large set of more specific traits or facets. Saucier and Ostendorf identified hierarchical subcomponents of the Big Five personality factors and they replicated the findings in English and German. In 1997, Saucier and Ostendorf found that 18

subcomponents were necessary to provide a finely faceted model for the Big Five. Based on their results, extraversion consists of four facets—sociability, unrestraint, assertiveness, and activity; agreeableness consists of four facets—warmth, gentleness, generosity, and modesty; conscientiousness consists of four facets—orderliness, decisiveness, reliability, and industriousness; emotional stability consists of three facets—low irritability, low insecurity, and low emotionality; and openness to experience consists of three facets—intellect, imagination, and perceptiveness.

3. CRITICISM OF THE TRAIT APPROACHES AND THE FIVE-FACTOR MODEL OF PERSONALITY

During the 1960s and 1970s, the trait approaches to personality were challenged based on a number of criticisms, mainly the lack of empirical evidence supporting the notion of temporal stability and cross-situational consistency. The failure to find evidence supporting these assumptions drove many personality researchers to conclude that traits were a theoretical construct useless for explaining personality and individual differences. In 1968, Mischel made a particularly relevant criticism of the trait approaches. He claimed that personality test scores (presumably reflecting traits) seldom correlated higher than .30 with other behavioral criteria. This was interpreted as strong evidence that behavior depends on situations.

The first trait approaches to personality structure were also criticized for other reasons. For example, Cattell's and Guilford's systems were criticized because their factors were not generally replicable across gender, age, and methods, and many researchers have failed to find them. Eysenck's P-E-N system was criticized on the basis that the psychoticism dimension is an amalgam of agreeableness and conscientiousness.

The FFM is currently the most influential model of personality; however, it is not unanimously accepted, and some criticisms have been made. For example, Block noted that the FFM is mainly based on factor analytical results, but in his view factor analysis is not an appropriate and sufficient base to decide the theoretical constructs of personality. Block also criticized the assumptions of the lexical hypothesis and the variability of conceptions about the Big Five. Other researchers have criticized the model on the basis of the need for many factors (e.g., Eysenck) or an insufficient number of factors (e.g., Hough). Such criticisms were answered by Costa and McCrae and Goldberg and Saucier, who noted that the Big Five were reproduced a great number of times, with different factor methods, by different researchers, with different instruments, and in different languages. Others factors were found only in isolated samples. Consequently, the model is not exhaustive, but it manages to support a more comprehensive explanation of various facets of human personality. Peabody and De Raad suggest that the Big Five personality structure cannot be as universal as some influential researchers have stated.

4. TRAITS, HERITABILITY, AND GENETICS

In recent years, many researchers have examined the genetic origin of human traits, and hundreds of papers have been devoted to this issue. The main conclusion of this research is that genes make a major contribution to individual differences in personality, especially when assessed by questionnaires. In the past 10 years, genetic research on personality has mainly focused on the Big Five personality dimensions, both in America and Europe. Loehlin summarized five large twin studies in which extraversion and emotional stability were examined. All the studies used self-report questionnaires that had not been developed for assessing the Big Five (with one exception). The measures of extraversion and emotional stability (neuroticism) were the most represented. The total sample comprised 24,000 pairs of twin. The findings showed heritability estimates of 49% for extraversion and 41% for emotional stability. Furthermore, the findings showed that the environmental factors were very important but that the environmental influence was mostly due to nonshared environmental effects rather than to the shared environmental influences. Shared environment is defined as environmental factors that contribute to twin and sibling similarity, whereas nonshared environment is defined as environment factors that do not contribute to sibling similarity.

The results of recent studies using questionnaires explicitly developed for assessing the Big Five personality dimensions (e.g., NEO-PI-R) confirm that the genetic influences on personality are strong. For example, using the NEO-PI questionnaire, in 1996 Jang and colleagues found that genetic influences accounted for

41, 53, 61, 41, and 44% for emotional stability, extraversion, openness, agreeableness, and conscientiousness, respectively. In 1998, Leohlin and colleagues, using 47 self-rating scales assessing the Big Five, estimated the genetic influence at 58, 57, 56, 51, and 52% for the same personality dimensions, respectively. In 1997, Riemann and colleagues, using the German self-report version of the NEO-FFI, found that genetic influences explained 52, 56, 53, 42, and 53% of the variance for the same dimensions, respectively. These studies examined all the variance sources using self-reported questionnaires, but there is also evidence from peer-reported questionnaires. In this case, the results suggest that genes account for two-thirds of the variance and nonshared environment accounts for the remaining one-third. Riemann and colleagues found that genes explained 61, 63, 81, 57, and 71% of the variance in the peer reports for emotional stability, extraversion, openness, agreeableness, and conscientiousness, respectively.

Additional information on the genetic and environmental sources of influence on personality was obtained by Borkenau *et al*. in 2001. These researchers independently examined the three sources of variance using ratings provided by judges who never met the twins but observed videotaped behaviors of one twin of each pair in 1 of 15 different settings. The aggregated video-based trait ratings, which were highly reliable (reliability mean = .94), showed that the genetic influences for the Big Five ranged from .38 to .62; that the shared environmental influences ranged from .20 to .28, with the exception of extraversion, which was not influenced; and that the nonshared environmental influences ranged from .30 to .38. In all cases, the genes were the main source of variance (Table I). An advance in clarifying the role of the genetic influences was made by Jan *et al*. in 2001. They found that the serotonin transport gene accounted for 10% of the relationship between emotional stability and agreeableness.

TABLE I
Genetic and Environmental Influences on the Big Five Personality Dimensions[a]

Dimension	*Genes*	*Shared Environment*	*Nonshared Environment*
Agreeableness	.43	.27	.30
Conscientiousness	.38	.25	.37
Emotional stability	.50	.20	.30
Extroversion	.62	.00	.38
Openness	.39	.28	.33

[a]*Source*. Borkenau *et al*. (2001).

5. THE BIG FIVE AND ORGANIZATIONAL BEHAVIOR

For many years, personality measures were considered poor predictors of organizational behaviors. This belief was mainly due to the small correlations found between personality characteristics and the organizational criteria (e.g., job performance) and in part due to the strong criticisms made by those with the situationist view of personality. However, during the past decade, many studies have shown that the Big Five personality dimensions are relevant variables for explaining many organizational behaviors and processes. Several organizational areas of interest for researchers were job performance, training success, teamwork, deviant behaviors, turnover, job satisfaction, leader emergence, leadership effectiveness, and transformational leadership behaviors. There are many cumulative results that allow to conclude that traits are relevant variables for predicting and explaining these organizational behaviors. For example, job performance is related to conscientiousness and emotional stability. Training success is mainly related to extraversion, openness, and conscientiousness. For teamwork, the dimensions with a strong relation are emotional stability, agreeableness, and conscientiousness. In the case of deviant behaviors, the best predictors of the absence of these dysfunctional behaviors are agreeableness and conscientiousness. For turnover, the most related dimensions are emotional stability, extraversion, agreeableness, and conscientiousness. Leadership behaviors were examined in relation to the emergence of leaders, the effectiveness of leaders, and the transformational leadership. In the three cases, the Big Five were shown to be relevant variables related to these organizational behaviors. For example, emotional stability, extroversion, openness, and conscientiousness were found to be related to the emergence of leadership. The effectiveness of leaders was associated with emotional stability, extraversion, openness, and agreeableness. Lastly, extraversion, openness, and agreeableness were related to transformational leadership behaviors (Table II). Thus, the dispositional variables of personality are thought to be relevant explanatory causes of many organizational behaviors and processes.

TABLE II
Correlations among the Big Five and Organizational Behaviors and Process[a]

Personality transformational dimension	*Job performance*	*Training success*	*Deviant behavior*	*Turnover*	*Teamwork*	*Leadership emergence*	*Leadership effectiveness*	*Leadership*
Emotional stability	.13	.08	–.06	–.35	.20	.24	.22	.03
Extraversion	.13	.23	.01	–.20	.16	.33	.24	.28
Openness	.13	.24	.14	–.14	.12	.24	.24	.26
Agreeableness	.07	.11	–.20	–.22	.27	.05	.21	.32
Conscientiousness	.31	.23	–.26	–.31	.23	.33	.16	–.06

[a]*Sources*. Barrick *et al.* (2001), Judge and Bono (2000), Judge *et al.* (2002), and Salgado (2002).

See Also the Following Articles

Agreeableness ■ Conscientiousness ■ Emotion ■ Extroversion–Introversion ■ Openness to Experience ■ Optimism ■ Personality Assessment ■ Psychological Assessment, Standards and Guidelines for

Further Reading

Barrick, M., Mount, M. K., & Judge, T. A. (2001). Personality and performance at the beginning of the new millennium: What do we know and where do we go next? *International Journal of Selection and Assessment*, *9*, 9–30.

Block, J. (1995). A contrarian view of the five-factor approach to personality description. *Psychological Bulletin*, *117*, 187–215.

Borkenau, P., Riemann, R., Angleitner, A., & Spinath, F. M. (2001). Genetic and environmental influences on observed personality: Evidence from the German observational study of adult twins. *Journal of Personality and Social Psychology*, *80*, 655–668.

Digman, J. M. (1990). Personality structure: Emergence of the five-factor model. *Annual Review of Psychology*, *41*, 417–440.

Goldberg, L. R. (1990). An alternative "description of personality": The Big Five factor structure. *Journal of Personality and Social Psychology*, *59*, 1216–1229.

Jan, K. J., Hu, S., Livesley, W. J., Angleitner, A., Reimann, R., Ando, J., Ono, Y., Vernon, P. A., & Hamer, D. H. (2001). Covariance structure of neuroticism and agreeableness: A twin and molecular genetic analysis of the role of the serotonin transporter gene. *Journal of Personality and Social Psychology*, *81*, 295–304.

Jan, K. L., Livesley, W. J., & Vernon, P. A. (1996). Heritability of the Big Five personality dimensions and their facts: A twin study. *Journal of Personality*, *64*, 577–591.

Judge, T. A., & Bono, J. E. (2000). Five-factor model of personality and transformational leadership. *Journal of Applied Psychology*, *85*, 751–765.

Judge, T. A., Bono, J. E., Ilies, R., & Gerhardt, M. W. (2002). Personality and leadership: A qualitative and quantitative review. *Journal of Applied Psychology*, *87*, 765–780.

Loehlin, J. C., McCrae, R. R., Costa, P. T., & John, O. P. (1998). Heritabilities of common and measure-specific components of the Big Five personality factors. *Journal of Research in Personality*, *32*, 431–453.

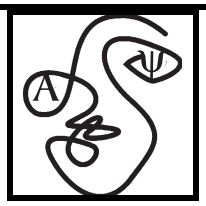
McCrae, R. R., & Costa, P. T., Jr. (1990). *Personality in adulthood*. New York: Guilford.

Peabody, D., & De Raad, B. (2002). The substantive nature of psycholexical personality factors: A comparison across languages. *Journal of Personality and Social Psychology*, *83*, 983–997.

Plomin, R., DeFries, J. C., McClearn, G. E., & Rutter, M. (1997). *Behavioral genetics* (3rd ed.). New York: Freeman.

Salgado, J. F. (2002). The Big Five personality dimensions and counterproductive behaviors. *International Journal of Selection and Assessment*, *10*, 117–125.

Saucier, G., & Ostendorf, F. (1999). Hierarchical subcomponents of the Big Five personality factors: A cross-language replication. *Journal of Personality and Social Psychology*, *76*, 613–627.



Transfer of Learning

Robert E. Haskell
University of New England, Old Orchard Beach, Maine, USA

1. Overview
2. Conceptual Issues and Problems
3. Categorical Distinctions
4. Models of Transfer
5. Historical Overview
6. Transfer Thinking and Reasoning
7. Higher Order Transfer
8. Instructional Issues
9. Dispositional Characteristics and Aspects of Transfer
10. Biological Origins of Transfer
11. Future Empirical Research and Theory
Further Reading

GLOSSARY

analogical reasoning/transfer Based on analogical proportionality (e.g., A : to B :: C : D).
causal A variable responsible for causing an effect or event.
counterintuitive Counter to what seems reasonable to expect.
induction A form of reasoning in which a small set of facts is used to justify a more general assertion.
isomorphic relations One-to-one structural relations between two phenomena that preserve the relations among their elements.
model A simplified working representation of a phenomenon.
structural relations Deeper underlying causal relations.
taxonomy A descriptive and systematic classification scheme.

Transfer of learning is considered to be the use of past learning in the learning of something new and the application of learning to both similar and new situations. The term transfer is derived from *trans*, meaning across or over, and *ferre*, meaning to bear, thus, to carry over.

1. OVERVIEW

Transfer of learning has a long history, beginning in instructional contexts and subsequently examined by laboratory research. Long considered a foundational educational concept, research on transfer of learning in formal educational settings has been less than encouraging. In addition, transfer of learning has been plagued by problems of definition and lack of coherent theory. In recent years, the concept has undergone reconceptualization from an educational and instructional concept to reflect a cognitive function undergirding thinking and reasoning. Because of its underlying basis in similarity relations, the concept of transfer of learning implies involvement in a host of other concepts, such as analogical reasoning, isomorphic relations, and generalization processes.

1.1. Conceptualizing Transfer

Beyond the general definition, although there is some disagreement regarding precisely how to define transfer, it is clear that it involves seemingly simple

©2004 Elsevier Inc.
All rights reserved.

similarity relations. The everyday basis of transfer can be exemplified by the following phrases: "it's like . . .," "it's equivalent to . . .," "it's akin to . . .," "for instance . . .," "it's the same as . . .," "similarly . . .," "in the same way . . .," "it resembles . . .," or "it's analogous to . . .." Accordingly, transfer as a cognitive function is involved in the use of figurative language as with analogies and metaphors. In addition, although not generally seen as involving transfer, because they are seen simply as an "instance" within a category, whenever an "example" is invoked transfer already has occurred. Examples are typically generated on the basis of similarity relations between the example and what it is seen as exemplifying. It can be seen that undergirding the cognitive concept of transfer and its various exemplifications is an invariance operation that may be possibly biologically based (Fig. 1).

1.2. Exemplifications

Transfer is pervasive in everyday activities. New things often are understood in terms of past experiences, as being like something already familiar. When first invented, cars were seen as "horseless carriages." With the advent of the computer age, there is now electronic "mail" and "notepads" as well as electronic " notebooks." Transfer is indicated when experience riding a bicycle is applied to learning to ride a motorcycle, when experience playing the piano is applied to playing an accordion, when experience ice skating is applied to roller blading, or when knowledge of one word processing program is applied to learning another. These examples are generally considered near transfer.

In formal educational settings, there is supposed to be transfer both within and among different knowledge domains and courses. For example, it is expected that there will be transfer or the application of what is learned from a mathematics course to everyday situations that require quantitative calculations.

On a higher level, transfer is manifested by recognizing that a concept in one subject area is like a concept in another seemingly unrelated subject area. For example, although it is seldom recognized, the concepts of learning in psychology (the process of acquiring knowledge or skill through practice, training, or experience), socialization in sociology (the social process whereby individuals acquire the values and behavior patterns of their society or social group), adaptation in biology (the alteration in an organism resulting largely from natural selection enabling the organism to survive), and acculturation in anthropology (the process of acquiring cultural traits or social patterns of another culture) are very similar. Recognition of these examples generally would be considered far transfer.

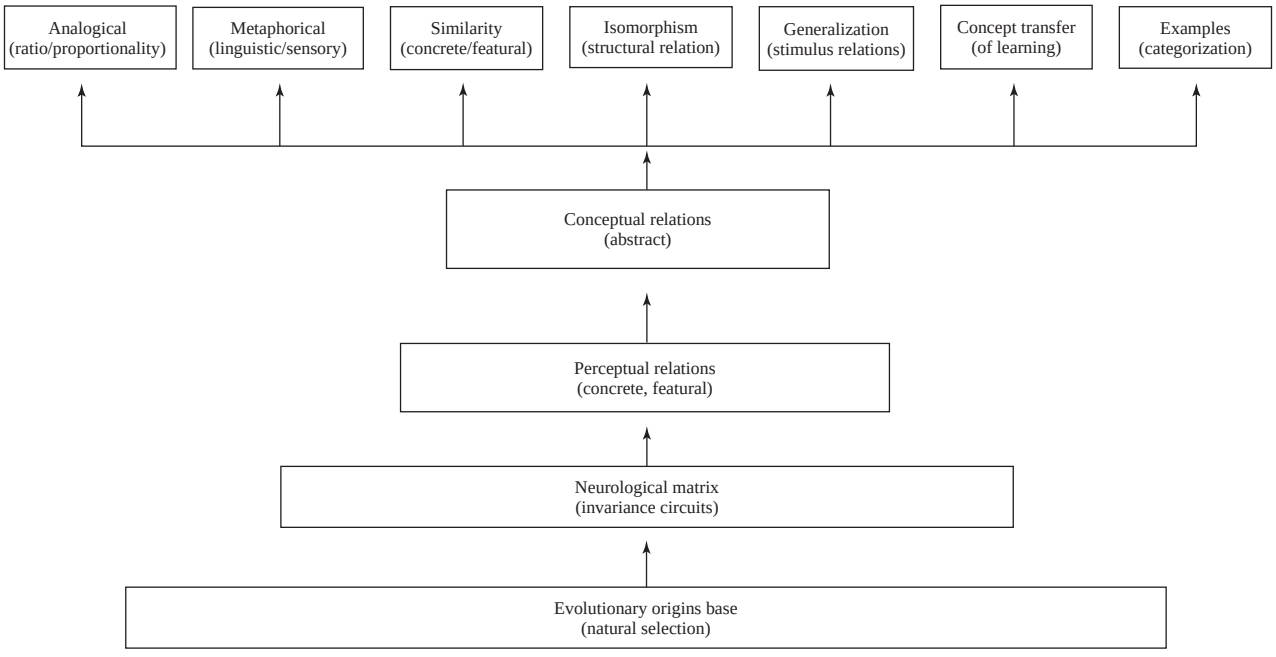


FIGURE 1 Transfer relations of invariance: domains and origins.

1.3. Importance

For more than 100 years, it has generally been agreed that the aim of all education—elementary, secondary, vocational, industrial, and business training—is to understand, extend, and apply what is learned in one context to similar and new situations. There continues to be a consensus that transfer is fundamental to all learning. Accordingly, transfer has become what Resnick characterized as the Holy Grail of education.

The demands of modern civilization make skill at transfer increasingly important. In a highly complex, rapidly changing, information-age society, the ability to transfer or generalize from the old to the new not only renders the world predictable and understandable but also is a necessity for adaptation to increasing technological and global advances. This is especially true where job and vocation changing are becoming the norm. Transfer is also seen by some as an important conceptual tool for interdisciplinary understanding and work.

1.4. Cognitive Benefits of Transfer

Although not typically stated in cognitive terms, transfer is extremely economical in terms of learning resources. For example, learning the aforementioned four concepts of learning, socialization, adaptation, and acculturation, each from a different discipline, requires a certain amount of memory storage space. Recognizing that the concepts are essentially the same, they can then be "chunked" into one larger concept, thus easing the load on memory storage. After the four concepts are seen as being alike, they become associated with each other, as well as other material associated with them, thus facilitating retrieval of the information. The four concepts then become cognitively integrated for a deeper understanding of each one. Finally, since the four concepts are not identical, certain aspects from one concept may reveal something new about the others that was not known prior to the transfer integration.

2. CONCEPTUAL ISSUES AND PROBLEMS

Despite the seemingly simple definition and examples of transfer given earlier, research reveals three basic problems of conceptualization involving not only the term but also its application, research design, and instructional outcomes.

2.1. Terminological Issue

There have been problems with the conceptualization of the term transfer, with some critics maintaining that the term is simply a metaphor since nothing actually gets transferred or "moved" from one domain or from one idea to another. The position is that all prior learning either simply enhances or, conversely, inhibits learning. In this view, the concept of transfer has no special merit or claim. Although the argument has some logical validity, if only for pragmatic and research purposes, most do not adopt this view. Nevertheless, the traditional distinction between the terms learning and transfer of learning is increasingly being questioned, with transfer being seen as primary.

2.2. The Question of Application Transfer

The second problem of conceptualizing transfer involves the expansion of its meaning. Although the extension of the original educational term to reflect a full-fledged cognitive function increasingly is accepted, it has been expanded in another way that does not have such acceptance. For example, when calculating the area of a 10×10 in. square learned in a math class, which then is applied to finding the area of an actual 10×10 ft. floor, it is generally considered to be a transfer of the classroom learning. Critics maintain that neither simple application of an abstract formula to a concrete situation nor the simple learning of finding the square feet of a classroom floor, when applied away from the physical place it was learned, should be considered transfer.

2.3. The Question of Facilitated Transfer

A third problem involves assessing whether transfer actually has taken place in a given instructional or research setting. Perhaps the leading critic of contemporary transfer research is Detterman, who argues that most transfer studies that demonstrate some low level of transfer in fact do not demonstrate transfer. For example, in one representative study children learned to stack toy tires on top of each other, eventually allowing a doll to reach a shelf on which other tires were to be stacked. Then, a similar problem was presented that required children to stack bales of hay on top of each other so that a farmer doll could reach a tractor. Three sets of such problems were used. It can

be seen that although the surface structure of each story in the set was slightly different, the underlying structure was transparently the same. If children were not able to solve a similar problem, the experimenter demonstrated the solution to them. Children were always asked to repeat the solution to the problem, at which time the next structurally identical problem was presented. Detterman maintains that the children did not engage in transfer but simply learned the "rules of the game" or were repeatedly shown the transfer over the three problem sets.

There are other methodological problems in studies that claim to find transfer, such as when subjects are told that previous material may be useful in the solution of a new problem, are explicitly informed about strategies and methods known to improve learning on the material to be learned, are instructed to use those strategies on that material, have a similarity of the material pointed out to them, and/or are given other hints about a similarity between the problems to be solved. Critics maintain that it is not reasonable to refer to the learning or problem solution as the result of transfer. The assumption behind such critiques is that only unaided or spontaneous transfer constitutes transfer. Based on research findings, however, spontaneous transfer seems to occur only 20% of the time in both adult and child populations (using age-appropriate tasks).

2.4. Outcome Research

Consistent findings during the past 100 years on which most researchers agree are that significant transfer in formal instructional settings seldom occurs, but it is pervasive in everyday situations; learning often is "welded" to the specific concrete context in which it is learned; when transfer does occur in formal instructional situations, it is frequently on such a basic conceptual level that many researchers do not consider it transfer at all but simple learning and application; and even when transfer occurs in informal everyday settings, it, too, is typically of a basic character scarcely distinguishable from simple learning or application of what is learned.

When more liberal assessments of the outcome literature are made, it is typically the result of less demanding conceptions or definitions of what constitutes transfer. It is not only general or significant (far) transfer that seems to occur infrequently but also the lower level application (near) transfer. It is widely agreed that the educational, social, technological, industrial, and business costs of transfer failure are considerable.

That transfer is pervasive in everyday settings is only seemingly contradictory to the negative findings in formal settings. If transfer is a general brain-based function, then it is reasonable to expect it to be pervasive in everyday settings on some level. However, the problem is to prevent transfer from being welded to the particular area in which it was learned, such as when transfer of mathematical operations occurs within mathematics but does not transfer to solving physics problems. In addition to the varying definitions, many variables have been identified that account for the failure of transfer, including insufficient learning of the original material, insufficient time encoding the original learning, inappropriate encoding of the material (e.g., rote vs deep or meaningful), insufficient learner motivation, inadequate memory span, the time lapse between original learning and the transfer task, insufficient learner background knowledge, and inappropriate instructional methods.

3. CATEGORICAL DISTINCTIONS

As with any set of phenomena, it is possible to classify transfer along various dimensions. Accordingly, there have been various general taxonomic schemes suggested, as well as other instructional distinctions.

3.1. Taxonomies of Transfer

Attempts to construct a systematic taxonomy or classification system of transfer continue to be limited in scope. This is due, in part, to the problem of precisely conceptualizing and quantitatively measuring what is called the metric of similarity involved in transfer (e.g., near versus far transfer). Nevertheless, any systematic, albeit qualitative, attempts can be useful for providing a common frame of reference or classification system for understanding and researching transfer, as the systematic scheme by Barnett and Ceci shows.

One traditional and continuing broad scheme is near versus far transfer. The former refers to transferring or applying what is learned to something closely similar; the latter refers to transferring what is learned to something quite dissimilar. The problem with such taxonomies is that they are relative not only to definition but also to individual perception of similarity on the part of both the learner and the evaluator, as well as to context variables that influence what is perceived as "near" and "far." What may be near transfer to one person may be far transfer to another. Transferring

within a domain of knowledge, however, is generally considered near transfer, and transferring across domains is considered far transfer.

In addition, there are many possible ways for classifying the different kinds of transfer, which include subject matter transfer, procedural or skill transfer, concrete transfer, abstract transfer, theoretical transfer, positive transfer, and negative transfer.

3.2. Positive and Negative Transfer

Transfer is considered positive when past experience leads to apprehending a valid or useful similarity relation. In general, negative transfer occurs when past experience leads to apprehending or applying an invalid similarity relation.

Negative transfer can also be the consequence of either insufficient knowledge or past learning interfering with current learning. An example of the former is when a whale is incorrectly thought to be a fish, or, conversely, when a porpoise and a deer are not seen as both being mammals. An example of the latter is when driving a vehicle with a clutch inappropriately activates the search for a clutch pedal the first time one drives a vehicle with an automatic transmission.

3.3. Transfer of Training versus Transfer of Learning

The long-standing distinction between learning or educating, on the one hand, and training, on the other hand, applies to instructing for transfer as well. Historically, the basic differences between training and learning are that training tends (i) to be more task specific, (ii) to be more concrete, (iii) to be less conceptual and theoretical, (iv) to be less process- and more immediate-outcome oriented, (v) to use a narrower range of knowledge base, (vi) to be less long-term outcome oriented, (vii) not to be as concerned with linkages to other knowledge bases, (viii) to result in a shallower depth of understanding, and (ix) to be less personally meaningful to the learner. Both training and educating are appropriate for their respective purposes.

Unlike in education, corporate instructional programs historically have been oriented toward training and the transfer of what has been learned in the training setting back to the workplace—a near application transfer based on physical distance. For both understanding and promoting transfer, the business model approach tends to focus on managerial and organizational support systems, whereas an academic or educational approach to promoting transfer focuses on instructional and cognitive processes. Both approaches seem to be needed to promote significant transfer in most instructional settings (for some business organizations, within approximately the past decade there has been a shift from an industrial-age training model to an information-age learning model) (Table I).

4. MODELS OF TRANSFER

Throughout the years, numerous models of transfer have been suggested. Early models once thought to be invalid continue to be examined and new models continue to be developed. The most well-known are described here. Some models are concerned with instructional processes, others with understanding the cognitive mechanisms of transfer, and still others with how to promote transfer through learning strategies.

TABLE I
Idealized Summary of Transfer Characteristics

Transfer of learning	*Transfer of training/basic learning*
Applying learning across domains/disciplines	Applying learning within a domain/discipline
Based on conceptual "elements"	Based on identical elements
Conceptual/underlying principle based	Concrete based
Context independent	Context dependent
Breadth and depth of scope	Narrow depth and scope
Generalization	Task specific
Figurative language/ metaphorical/analogical	Literal language/domain examples
Larger and general knowledge base	Restricted to task knowledge base
Theory based	Immediate information based
Process oriented	Product oriented
Dispositional based	Instrumental/task based
Higher order/progressive transfer	Single-level transfer
Far transfer	Near transfer

4.1. Formal Discipline Model

This model developed out of ancient and classical educational theory, in which it was thought that the internal logic or structure of a discipline such as geometry, mathematics, or Latin somehow enabled a learner to transfer automatically what was learned in these disciplines to other areas. For example, it was thought that courses in formal logic taught people how to think in everyday situations. From early on, little rigorous research supported the formal discipline approach to transfer. Currently, with some interesting exceptions, the formal discipline approach is considered antiquated. However, at least as early as 1937, Allport, the renowned psychologist, suggested that the formal discipline model should not be dismissed entirely, and later, Ellis also suggested that the formal discipline model may not be a dead issue but simply in need of reformulation. Recent research suggests that the formal discipline model may have validity in some domains (e.g., statistical reasoning).

4.2. Behavioristic Models

4.2.1. Identical Elements Model

The identical elements model, developed from the classic laboratory research of Thorndike and Woodworth in 1901, maintains that for transfer to occur there must be a common set of concrete identical elements between two experiences or learning sets. The identical elements model research resulted in the initial demise of the formal discipline model of transfer. Thorndike's negative findings of transfer outcomes led to a pessimistic view of the possibility of teaching for transfer and has exerted wide-ranging influence over instructional and curriculum design until present day.

A neo-identical elements model remains extant in some areas of cognitive psychology research. For example, what are called production rules (essentially "if . . . then" statements) in computational models of cognition are considered higher level versions of the identical elements stimulus–response connections of Thorndike. Much of the research on analogical reasoning (transfer) is also based on identical element relations of some kind, either concrete features or abstract relations.

4.2.2. Stimulus Generalization Model

Almost from its inception, transfer research has been associated with the concept of stimulus generalization developed from the behaviorist laboratory research of Pavlov's classical conditioning paradigm. Stimulus generalization is the evocation of a nonreinforced response to a stimulus that is very similar to an original conditioned stimulus. The stimulus generalization model presupposes identical elements that enable generalization to occur. Stimulus generalization, then, can be viewed as a basic physiological learning "theory" explanation of how transfer occurs. Despite this historical association with generalization, virtually no recent empirical or theoretical work has been conducted.

4.3. Cognitive Models

4.3.1. General Principle Model

Just as Thorndike challenged the formal discipline model, the classic research of Judd in 1908 challenged the identical elements model by findings that suggested transfer also occurs on the basis of learners' understanding the abstract general principle(s) underlying phenomena that then can be applied to situations that do not possess obvious identical elements but that have the same or similar underlying principle. Judd's model can be considered a precursor of some current cognitive approaches to transfer because the general principle model is concerned with what is in learners' heads, not just the featural similarity between stimuli or within instructional material. It also can be considered a precursor to the metacognition approach because it recognizes that transfer can be promoted by strategies such as directing a subject's "attention" to the general principles involved.

4.3.2. Information Processing Model

Historically, the concept of transfer of learning lay squarely within the applied instructional arena of education. Accordingly, until recent decades, cognitive psychology developed no theory of transfer per se. Currently, transfer as a function figures prominently in many areas of cognitive research. With few exceptions, this research has not been concerned primarily with instructional processes but rather with laying the foundation for an information processing framework that explains the internal representation, processing, and retention of information. One central area in contemporary cognitive psychology directly relating to transfer is schema research. As Phye notes, the information processing model tends to emphasize learner performance, not learner acquisition of knowledge (i.e., learning).

4.3.3. The Metacognition Model

This model is concerned primarily with promoting transfer, and not explaining its cognitive dynamics. The metacognition model makes use of self-monitoring strategies, leading learners to transfer these strategies within and across tasks or learning domains. Strategies include learners' self-monitoring to observe what they are doing or thinking while they are learning. Metacognitive research is viewed by many as the "new approach" to transfer. Strictly speaking, the metacognitive approach is not a part of cognitive psychology research but is more aligned with education.

5. HISTORICAL OVERVIEW

As noted previously, experimental research on transfer began early in the 20th century with psychologists Thorndike and Woodworth (1901) and Judd (1908); the first two are identified with the identical elements model of transfer and the latter with the general principle model. These two opposing models can be seen as progenitors that continue to frame much of the modern research agenda on transfer.

More than 100 years of transfer history shows that until the past few decades, (i) the identical elements model overwhelmingly predominated educational research and instruction and remains a strong influence, although to a less significant degree; (ii) in textbooks, transfer was given only cursory mention; (iii) the educational field had ceased to generate useful transfer findings or applications; (iv) transfer chapters largely had been missing from the pages of educational textbooks; (v) there was no comprehensive educational text on transfer; (vi) the concept of transfer had been included in neither cognitive psychology research nor cognitive psychology texts; and (vii) there had been an almost complete absence of conceptual and theoretical development of transfer as an educational construct.

5.1. Research Publications

According to Clark and Voogel's review of the educational literature, from the 1960s to the mid-1980s, research on transfer per se seemingly declined. Why this decline occurred had partly to do with the recognition that the identical elements model of transfer had ceased to be further developed and a shift in instructional focus to specific skills and competencies.

This apparent decline was misleading, however, because other fields had begun to reconceptualize transfer and to relabel it. For example, basic cognitive schemata have come to be seen by some as involving transfer. Because of an increasing recognition of the cognitive operations involving similarity relations, research on transfer has been extended to include the growing cognitive research on metaphorical and analogical transfer as well as what are called isomorphic relations, in which stories of identical structure but different content are used to test for transfer. Other research includes the cognitive function of using a variety of examples in instructing for transfer. Because cognitive research tended to take an information processing approach, which focused on how knowledge is presented and processed rather than on how knowledge is acquired in the educational or learning sense, the research was not concerned with the educational and instructional aspects of knowledge acquisition and transfer.

Historically, numerous edited books compiling articles on transfer were available, but there was no comprehensive nonedited volume. It was not until 1965 that the partially edited book by Ellis was published. It was the first attempt at a comprehensive integration of transfer research. In 1989, Singley and Anderson published a specialized nonedited computational approach to understanding transfer of skills. These books were concerned only marginally with instruction for transfer. In 1995, McKeough *et al.*'s edited volume was focused specifically on instruction for transfer based on the current research. It was not until 2000, however, that the first nonedited comprehensive overview of the research from an educational and cognitive perspective, with an integrated framework on instruction and theory, was published by Haskell.

5.2. Conceptual and Instructional Shifts

With few exceptions, most of the research attempting to discover the underlying mechanisms of transfer had been in cognitive psychology and not in educational research. Thus, the term transfer of learning, the original superordinate and applied educational construct, had been in danger of becoming a fragmented and subordinate process subsumed under the cognitive research labels outlined previously. Although it was repeatedly noted that educators should be exhorted simply to "teach for transfer," it was neither sufficient nor effective instructionally. Until recent decades, as Royer notes, with few exceptions

the educational field did not significantly adapt cognitive research findings to instruction.

In the early 1980s, educational researchers began to be aware of and to adapt the growing cognitive research related to transfer. This shift to a cognitive perspective is thought by some to have had a double edge, however. On the positive side, it shifted emphasis from a cognitively uninformed instructional view of transfer to a more researched-based construct, including the cognitive research on learning strategies that are thought to promote transfer. On the negative side, pure research into the cognitive mechanisms underlying transfer tended to overshadow the applied instructional approach to transfer by focusing on understanding the cognitive mechanisms of transfer. With the exception of the metacognition approach, this shift of focus from transfer as an applied instructional concept to one emphasizing a cognitive information processing approach constituted a kind of paradigm shift that is still quite evident.

6. TRANSFER THINKING AND REASONING

In contrast to the traditional view of thinking and reasoning, it is becoming increasingly accepted that most everyday reasoning does not involve the manipulation of abstract rules or otherwise symbolic systems resembling a formal logic. Rather, it seems to involve pragmatic reasoning schemata, mental models, and analogical reasoning; hence, it involves transfer. Much research has demonstrated the pervasiveness and importance of analogical transfer in thinking and reasoning. The considerable research on similarity clearly demonstrates its systemic role in everyday reasoning processes. Researchers of analogical and metaphorical reasoning suggest that the perception and manipulation of similarity is the foundation of thinking and reasoning. Indeed, both Plato and Aristotle recognized this.

6.1. Single-Instance Reasoning and Analogical Transfer

Research on the use of single instances in reasoning and decision making shows that people often place as much, if not more, credence in a single instance than in theory or logic. It is increasingly accepted that much of human reasoning is carried out by using schemata from everyday life as analogs that are tied to particular bodies of knowledge.

6.2. Legal Reasoning and Transfer

The use of single-instance reasoning has implications well beyond its use by the layperson. A variant on single-instance and analogical reasoning is legal reasoning. The finding of similarity is the key ingredient in legal reasoning. Thus, much of legal reasoning is based on transfer. As the history of legal reasoning shows, it is based largely on reasoning from a single or a few instances or cases. Attorneys and judges often argue whether some previous case is applicable to a current case. Even reasoning on the basis of legal principle involves transfer reasoning; judges often decide cases "on principle," demanding that like cases should be treated alike.

6.3. Social Policy and Decision Making

A number of studies on analogical transfer have noted that historians and governmental policymakers often make use of single instances in their experience as analogs on which to base decisions and policies about current crises. For example, in 1950, North Korea invaded South Korea. Although South Korea was thought to be of no significant strategic value, President Harry S. Truman was intrigued by certain historical incidents that had preceded World War II that he considered analogous to the North Korean invasion, particularly the Japanese invasion of Manchuria, the Italian attack on Ethiopia, and the annexation of Austria by Germany. Truman saw that the invasion of South Korea was just the beginning of a series of hostile acts like those that led to World War II, just as President George Bush, Sr., saw Saddam Hussein's invasion of Kuwait as reminiscent of Hitler's actions.

6.4. The Validity of Transfer Thinking

Although widely and inescapably used in reasoning, transfer thinking should be approached cautiously. For centuries, philosophers have pointed out the dangers of reasoning based on similarity or analogy. The problem engendered by such reasoning can be seen as subject to the same problems as logical inference and induction. Recently, however, improved methods for assessing analogical transfer have been developed that apply to transfer thinking and reasoning.

7. HIGHER ORDER TRANSFER

Most applications of transfer are, in effect, analogical in nature, with some becoming quite complex in terms of abstract relationships and proportionality ratios. Transfer thinking is involved in even higher order thought sequences.

7.1. Transposition Transfer

Another phenomenon that can be seen as involving transfer is called transposition. Perceptually, for example, when animals conditioned to respond only to the darker of two gray squares are later shown an even darker gray square along with the original darker square, the animals respond to the new darkest gray square to which they had not been conditioned. This response is generally considered to occur because in relation to the first set of squares, the new darker square is perceived as the same as the original darker square. The transfer was not to the absolute specific shading, but the relative shading of the new darker square in relation to the shading of the original darker square. Notated analogically, the relation is as follows: the first darker square : the first lighter square :: the new darker square : the original darker square. Auditorially, transposition transfer can be illustrated by recognizing a musical piece played in a different key or octave as being the same even though no note is the same.

7.2. Progressive Transfer

Modeled on analogical and arithmetical progression, Haskell suggested a type of transfer not generally recognized that can be referred to as progressive or continuous transfer. This type of transfer takes the abstract form of arithmetical progression: 2, 4, 8, 16, 32. Stated in analogical form, it is 2 : 4 :: 4 : 8 :: 8 : 16 :: 16 : 32. As a semantic example, the biological classification system can be put into this progressive form as follows: species : genus :: genus : family :: family: order :: order : class, etc. Related to analogical reasoning, Aristotle referred to this form of reasoning as continuous analogy.

7.3. Transfer in Science

The history of science and invention is replete with those who are skilled at transfer. For example, Darwin, who developed human evolutionary theory, transferred the idea of farmers' selective breeding (i.e., artificial) of animals to his development of the principle of natural selection, in which natural processes were selecting genetic traits in place of the farmer. Physicist Louis De Broglie noticed that the mathematical equations of another well-known physicist, Neils Bohr, who described the orbits of an electron, were the same equations used to describe the vibrating waves of a violin string. With this transfer, De Broglie revolutionized atomic physics and laid the foundations of quantum mechanics.

As these two examples show, many advances in science are made on the basis of what, after the fact, seems to be a simple "it's like ..." type of transfer. These examples are included in this section because theoretical thinking and a complex knowledge base were required before the higher order of transfers could be made.

8. INSTRUCTIONAL ISSUES

Although all educational issues have an impact on transfer, arguably there are relatively transfer-specific ones that stand out in terms of achieving transfer, including the counterintuitive nature of instructing for transfer, learning strategies, knowledge base, contexts, practice, and variability.

8.1. Counterintuitive Nature of Transfer

A generally unrecognized problem in understanding, achieving, and instructing for transfer is that doing so is often counterintuitive to what might be expected from instructing general learning. For example, it is generally assumed that those with a high level of theoretical knowledge of a subject are more proficient at transfer than people with only concrete knowledge. Although this is typically true, there are important counterintuitive exceptions.

For example, students who use a deep learning or conceptual approach to material, as expected, tend to score higher on more complex recognition tasks but, unexpectedly, score lower on simple recognition tasks. Moreover, students who lack prior knowledge about a skill tend to do better than those who have prior knowledge of the skill—but only under rote learning conditions. Students who learn reading material in a

verbatim or rote fashion do better than students who learn the material in a conceptual fashion on verbatim tests, but students who learn the material conceptually are better at applying their knowledge in a more generalized way.

The point is that although conceptual or theoretical knowledge is advantageous for engaging in far transfer, it may be disadvantageous for simple concrete application situations. Such counterintuitive transfer outcomes across other subject areas yield similar findings, suggesting that if the goal of instruction is transfer of explicitly taught skills, then learning instrumental and concrete procedures may be more effective than conceptual learning. However, if the goal of instruction is far transfer, then learning conceptually is superior to simple instrumental learning.

Counterintuitive findings also apply to conditions of practicing a skill. For example, Schmidt and Bjork found that providing information to learners about their performance errors during practice improves performance during training but can degrade performance on a test of long-term retention or transfer. Furthermore, increasing the task variability during practice decreases performance during training but facilitates performance on later tests of the ability to generalize training to new conditions.

8.2. Contexts and Cultures of Transfer

Typically, learning has been thought of as the inherent property of the individual learner. However, it is increasingly recognized that all learning takes place in a context and that the context typically becomes a part of, and associated with, the learning and its transfer. Broadly construed, the concept of context includes not only the immediate learning situation but also sociocultural influences. More narrowly, context differs from cultural influences in that context tends to be associated with the learning situation. What Pea has called cultures of transfer, however, are networks of social norms that may either enhance or inhibit transfer.

8.3. Knowledge Base, Learning Strategies, Practice, Theory, and Exemplars

After a long period of concentrating on a content-empty learning strategies approach to instruction, it is increasingly recognized that content-specific knowledge (or declarative knowledge) is crucial for transfer. An appropriately encoded knowledge base results in expertise, and expertise is associated with superior transfer. Theoretical knowledge is also central to achieving transfer. This is true even for small children. Historically, it has been thought that young children are "perceptually bound" to the concrete features of stimuli and are unable to transfer causal relations appropriately. However, the seminal work of Brown demonstrated that young children can engage in transfer generated from a "deep structure" level of conceptual understanding even when similar surface elements between tasks contraindicate transfer.

In addition, with some exceptions, research strongly suggests that reflective practice facilitates significant transfer. The research also shows that, with only minor exceptions, one of the most important means of facilitating transfer is with the use of a variety of related examples. It is thought that processing multiple and varied examples results in the development of cognitive schemata.

After reviewing the research on transfer, Haskell suggests that for significant learning and transfer to occur, the following principles are required: Learners must (i) acquire a large primary knowledge base in the area in which transfer is required, (ii) acquire some level of knowledge base on subjects outside the primary area, (iii) understand what transfer of learning is and how it works, (iv) understand the history of the area(s) in which the transfer is wanted, (v) acquire dispositional characteristics, (vi) develop an orientation to think and encode learning in transfer terms, (vii) create contexts and cultures that support transfer, (viii) understand the theory underlying the area(s) they want to transfer, (ix) engage in hours of practice and drill, (x) provide time for the learning to incubate, and (xi) observe and read the works of those who are exemplars of transfer thinking.

9. DISPOSITIONAL CHARACTERISTICS AND ASPECTS OF TRANSFER

Dispositional aspects of transfer relate to personality and positive attitudinal characteristics toward learning, as Bereiter suggests. Strictly speaking, the dispositional view can be considered another model of transfer.

9.1. Personality Structure and Development

Although not widely recognized or researched, as early as 1937 Allport suggested that transfer may be important for the theoretical psychology of personality. Essentially, the idea is that for most people, the complex of behaviors and ideas that make up who they are remain unorganized, unrelated, and dissociated. Presumably, the ability at transfer may be used to integrate this complex of traits. Ausubel suggested that there exists two broad cognitive styles, generalizers and particularizers, with the former engaging in more transfer than the latter.

9.2. Disposition for Learning

A major issue in how best to achieve significant transfer is whether to instruct primarily for specific skills or to instill in a learner a personal disposition toward learning and transfer. The former involves acquiring abstract learning strategies, including metacognitive skills; the latter involves instilling motivational and affective values or a positive temperament for learning and transfer. As judged from the direction of most research, the general consensus seems to be with the former.

9.3. Prejudice

Transfer is also involved in both the development and the reduction of prejudice toward others. Prejudice can be seen as inappropriate or negative transfer of perceptions and experiences regarding others. Assuming positive transfer, this involves a person's inferring that his or her experience is similar in some degree to another person's experience, thus leading to better understanding of others. Accordingly, analogical transfers of experience can be useful for both understanding prejudice and understanding those with differences.

10. BIOLOGICAL ORIGINS OF TRANSFER

A possibility that has been largely overlooked is that transfer may have evolutionary origins and a neurological basis. For a phenomenon as pervasive and significant as transfer, it might be expected that it would have biological origins. Although the research is inconclusive, Haskell suggests that transfer has both evolutionary and neurological bases.

10.1. Evolutionary and Neurological

The evolutionary argument is that being attuned to even superficial or surface similarities may have bestowed a survival advantage on animals in complex environments. Surface similarities are swift evaluations that are often needed to avoid danger. In evolutionary terms, if it looks like a hungry lion and growls like a hungry lion, then there are survival advantages to assuming it is a hungry lion. In other words, it may have had more survival value to assume that a surface similarity is meaningful than to assume it is not. For more complex thinking, it is thought that surface similarity is often positively correlated with deeper, more "causal" or structural relations among phenomena. Evolutionarily, the consequences of an invalid surface similarity typically are not as serious as ignoring a valid one. Accordingly, a neurological substrate likely evolved. Haskell suggests findings indicating neurological circuits that specialize in the generalization of learned experience to new situations via invariance-generating circuits.

10.2. Transfer and Intelligence

Perhaps the most controversial issue regarding significant or far transfer is its relationship to innate intelligence, as measured by IQ tests. Historically, transfer has been thought to be correlated with IQ mainly because the lower the IQ of a learner, the less transfer is typically achieved. Ability at analogical transfer, as measured by the Miller Analogy Test, has long been correlated with IQ and, some suggest, with creativity. However the issue resolves, it is clear that IQ, although seemingly a necessary condition, is not a sufficient condition for transfer.

11. FUTURE EMPIRICAL RESEARCH AND THEORY

Future research on transfer will include most of the areas mentioned earlier, with specific attention to (i) how to instruct effectively for transfer; (ii) understanding how contexts influence transfer; (iii) individual learner characteristics; (iv) establishing a more precise taxonomy or classification system; (v) specific research on the various kinds of transfer, especially progressive; (vi) delineating the concept as a general cognitive function; and (vii) identifying the neurological substrates of transfer. Finally, an emphasis on theory construction is needed to guide empirical research and the development of the concept of transfer.

See Also the Following Articles

Educational Achievement and Culture ■ Learning ■ Learning Styles and Approaches to Studying ■ Teaching Effectiveness

Further Reading

Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn? A taxonomy for far transfer. *Psychological Bulletin*, *128*(4), 612–637.

Bereiter, C. (1995). A dispositional view of transfer. In A. McKeough, J. Lupart, & A. Marini (Eds.), *Teaching for transfer: Fostering generalization in learning* (pp. 21–34). Mahwah, NJ: Erlbaum.

Brown, A. L., & Campione, J. C. (1984). Three faces of transfer: Implications to early competence, individual differences, and instruction. *Advances in Developmental Psychology*, *3*, 143–192.

Clark, R. E., & Voogel, A. (1985). Transfer of training principles for instructional design. *Education Communication and Technology*, *33*(2), 113–123.

Cormier, S. M., & Hagman, J. D. (Eds.). (1987). *Transfer of learning contemporary research and application*. New York: Academic Press.

Detterman, D. K., & Sternberg, R. J. (Eds.). (1993). *Transfer on trial: Intelligence, cognition, and instruction*. Norwood, NJ: Ablex.

Gladstone, R. (1989). Teaching for transfer versus formal discipline. *American Psychologist*, *44*(8), 1159.

Haskell, R. E. (1998). *Reengineering corporate training: Intellectual capital and transfer of learning*. Westport, CT: Quorum (Greenwood Publishing Group).

Haskell, R. E. (2000). *Transfer of learning: Cognition, instruction, and reasoning*. San Diego: Academic Press.

Haskell, R. E. (2002). Cognitive science and the origin of lexical metaphor: A neurofunctional shift (NFS) hypothesis. *Theoria et Historia Scientarium*, *6*, 291–326.

Holyoak, K. J., & Koh, K. (1987). Surface and structural similarity in analogical transfer. *Memory and Cognition*, *15*(4), 332–340.

Pea, R. D. (1987). Socializing the knowledge transfer problem. *International Journal of Educational Research*, *11*, 639–663.

Phye, G. D. (2001). Problem-solving instruction and problem-solving transfer: The correspondence issue. *Journal of Educational Psychology*, *93*(3), 571–578.

Read, S. J., & Cesa, I. L. (1991). This reminds me of the time when . . .: Expectation failures in reminding and explanation. *Journal of Experimental Social Psychology*, *27*, 1–25.

Resnick, L. B. (Ed.). (1989). *Knowing, learning and instruction: Essays in honor of Robert Glaser*. Hillsdale, NJ: Erlbaum.

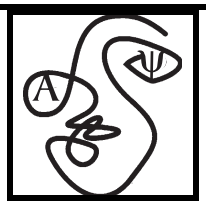
Schmidt, R. A., & Bjork, R. A. (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psychological Science*, *3*, 207–217.

Singley, M. K., & Anderson, J. R. (1989). *The transfer of cognitive skill*. Cambridge, MA: Harvard University Press.

Spellman, B. A., & Holyoak, K. J. (1992). If Saddam is Hitler then who is George Bush? Analogical mapping between systems of social roles. *Journal of Personality and Social Psychology*, *62*(6), 913–933.

Sunstein, C. R. (1993). On analogical reasoning. *Harvard Law Review*, *106*, 741–791.

Vogniadou, S., & Ortony, A. (Eds.). (1989). *Similarity and analogical reasoning*. Hillsdale, NJ: Erlbaum.



Translation

Richard W. Brislin, Brent MacNab, and David Bechtold

University of Hawaii, Honolulu, Hawaii, USA

1. Translation and Cross-Cultural Research
2. Back-Translation
3. Pretesting
4. Emics and Etics
5. Decentering
6. Writing Translatable English
7. Recent Enhancements of Translation Methodology in Research
8. Translation and Qualitative Studies
9. New Challenges in Translation: The Pursuit of Equivalency in Higher Order Issues
10. Future Directions of Translation and Translation Models
11. Conclusion

Further Reading

GLOSSARY

asymmetrical translations A literal translation of a research instrument that is used to identify differences between cultures.

back-translation Using independent bilingual translators, the act of translating material from host language to target language and then translating back into host source to determine if equivalence of meaning is achieved.

cultural emic A concept, value, or other artifact of culture that has particular meaning to a specific culture (e.g., dating).

cultural etic A concept, value, or other artifact of culture that has universal meaning across all cultures (e.g., love).

culture shock A sense of loss that accompanies leaving one's own culture and adjusting to another.

decentering A method of translation–back translation that allows for the development of a research instrument that is of equivalent meaning to both cultures under study.

host version The original language used in the version of material to be translated.

item response theory A method of comparing responses from different translated versions of questionnaires to determine if there is equivalence.

pretesting The use of a translated version in a pilot test to determine if the instrument and response are as expected and equivalent to the instrument and responses in the original version of the instrument.

symmetrical translations A conceptual translation of a research instrument that is used to identify similarities across cultures.

synchronous automated translation systems (SATS) Technology that can produce seamless translations with minimal time delay.

target version The language used in which the original version of the material is translated into.

Translation to achieve equivalence of meaning has become a significant issue for international and cross-cultural research. A failure to achieve equivalence in translation can put into question the research generally and conclusions derived from that research specifically. Beginning in the 1970s, efforts began to enhance the quality of translation in research through the use of various methods to achieve equivalence. The use of back-translation and decentering techniques were the first attempts at achieving meaning equivalence in translation. Recently, the use of dual-focus methods in

©2004 Elsevier Inc.
All rights reserved.

questionnaire development as well as quantitative strategies such as item response theory have been developed to further enhance translation effectiveness. Computer and technology are also becoming important tools in translation, with the development of software such as synchronous automated translation systems. These new tools are expected to enhance both speed and accuracy of translating materials from one language to another.

1. TRANSLATION AND CROSS-CULTURAL RESEARCH

In almost all cross-cultural research studies, methodological issues associated with translation between languages are certain to arise. Researchers have tests, questionnaires, and interview schedules, but they also have experimental protocols so that research participants know what is expected as they attempt to complete various tasks and activities. Translation issues are clearest when languages from very different linguistic families are involved, such as English and Japanese or Swedish and Chinese. However, it is also wise for researchers to know about translation issues even when they deal with dialects, such as English spoken in the United States compared with that spoken in Scotland or French spoken in France compared with that spoken in Canada.

Many early cross-cultural studies (i.e., during the 1950s and 1960s) have standardized tests as their starting point. Well-known tests, such as the Minnesota Multiphasic Personality Inventory or the Stanford–Binet Intelligence Test, were translated into other languages, administered in various countries, and interpretations made based on patterns of differences. Many of these studies are embarrassing by today's methodological standard and have been relegated to the deserved obscurity of dust-gathering journals on library shelves. Prior to 1970, there were few attempts to assess the quality of translations. Bilinguals were presented with a standardized test in one of the languages that they knew and wrote a version in their other language. Few researchers asked about issues such as the quality of the translation, equivalence of terms, familiarity of the words in the translated version, and the level of literacy needed among research participants to give meaningful responses to test items. The lack of attention to translation quality led to the publication of many problematic studies that are rarely cited today.

Even with these difficulties, lessons were learned that have proven useful in the development of cross-cultural psychology. One important lesson is that translation between languages is not a simple and direct process in which researchers can present wording in one language and expect straightforward equivalents in another language. Researchers in the 1950s and 1960s often expected such direct equivalence and had difficulties when they did not find it. Instead of direct equivalence, they found that familiar terms in their own language were difficult to translate. The importance of "punctuality" for workers was difficult to translate into languages where cultural practices do not include organizing the workday around clocks. "Friend" in American English was difficult to translate into languages where close nonfamily members are viewed as deserving of long-term obligations and collective memberships. Americans often use "friend" to describe a person they met and chatted with yesterday.

The realization that frequent nonequivalence is a common occurrence in cross-cultural research led researchers to deal with this fact in creative ways that ultimately advanced the quality and relevance of cross-cultural studies. For example, it led to the development of procedures such as the antecedent–consequent method, in which the meaning of terms such as friend could be probed. Questions of respondents in different cultures would take forms such as the following:

(Antecedent)
If you have people who ___, then they are friends.
If people ___, then they may become friends.
(Consequent)
If people are friends, they are expected ___.
If a person has problems, a friend will ___.

Responses to such inquiries will give insights into such issues as the length of time people know each other before becoming friends, activities in which friends engage, and the obligations that friendship entails. This information that surrounds a term such as friend, with rich context, is extremely useful for translators who must search other languages for the best equivalent term or phrase.

This realization concerning translation and change in attitude paralleled another change in cross-cultural studies that occurred at approximately the same time and also provided an additional term for translation analysis. Culture shock was a term first coined in 1958 by Kalvero Oberg. So persuasive did the term become that training professionals whose jobs included preparing people for life in other cultures quickly

incorporated it into their training programs. The term refers to the sense of loss and corresponding frustrations, emotional upheavals, and physiological changes that accompany leaving one's own culture and adjusting to another. Beginning in the 1960s, there was a large group of highly educated and articulate people who learned about culture shock in cross-cultural training programs. These were Peace Corps volunteers. Culture shock was at first treated as a problematic state that might mark unsuccessful volunteers. It was therefore introduced as an indicator of adjustment problems. This treatment of the term changed when the first waves of volunteers returned after their 2-year overseas assignments. Most had experienced culture shock but also felt that their assignments had been successful. Because most experienced culture shock, it was difficult to view it as a sign of maladjustment. The view of culture shock changed and (beginning in the mid- to late 1960s) it became viewed as an expected part of any overseas assignment. Translations of the term culture shock into other languages had to include this changed connotation.

2. BACK-TRANSLATION

Working to improve the quality of translation for research purposes, Werner and Campbell published an influential analysis of back-translation in 1970. With this method, monolingual researchers can gain some, although not perfect, insights into the quality of the translation by working carefully with bilingual translators. Back-translation is based on a set of steps. For ease of discussion, examples of translations from English are used (Fig. 1).

Although not perfect, this back-translation procedure provides a good starting point. The researcher can examine the two English language versions and if they are equivalent and if the back-translation is worded smoothly, then he or she is at "step one" of working toward a high-quality translation. Researchers should work closely with translators and treat them like respected colleagues. Researchers can ask the first bilingual, "Was there any English language material that was difficult to translate?" They can ask the second bilingual, "Was there any material in the target language that you had difficulty with, for example, finding equivalents in English." If translators know details about the goals of a research project, they can be extremely helpful in suggesting alternative wordings and sometimes indicating ideas for new research directions. For example, assume that the original English wording asks about the respondents' relationships with their brothers. A Japanese bilingual would point out that there is no equivalent for brother. In Japanese, native speakers must make a distinction between "older brother" and "younger brother." There are distinct terms based on age distinction but no general term for brother. If not integrated into the research team, a translator might choose one of the terms as a "default option," but this could interfere with the researchers' goals. This feedback from the translator might stimulate the researchers to examine the importance of status distinctions in Japan. They would find, for instance, that age distinctions are taken seriously even at a 4-year college. Sophomores take on a *sempai* role, or the role of superior and guide, to freshman *kohai*, or neophytes.

Why is back-translation only a good first step? One reason is that certain words have near-automatic translation equivalents, but the terms may not be conceptually equivalent, as indicated by a more detailed analysis such as the antecedent–consequent method. The English "friend" will translate to "amigo," and this will translate back to the English "friend" with near certainty. However, the types of people that respondents call friend and amigo may be very different. Similarly, the English "party" may translate into the French "soiree," and this will translate back into English as "party," but are the same types of social gatherings being discussed?

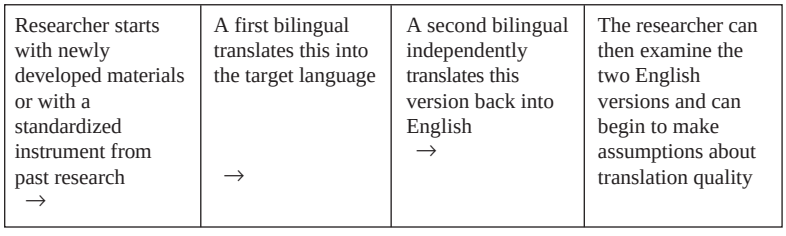

Researcher starts with newly developed materials or with a standardized instrument from past research
→
A first bilingual translates this into the target language
→
A second bilingual independently translates this version back into English
→
The researcher can then examine the two English versions and can begin to make assumptions about translation quality

FIGURE 1 Translation quality: translation and back-translation procedure.

Another reason for back-translation's limitations is that the second bilingual, if skillful, can make sense out of a poorly worded translation and then render it into very fluent English. The researcher will read the back-translated English and might conclude that the target version must be good given that the second English version reads so well. The activities of the second bilingual are similar to those done by many readers of this article. Haven't many of us been in social situations in which we tried to help international visitors or international students? They spoke to us in broken English, but we were able to figure out what they wanted and could express their wishes in fluent English: "You want the address of a good hotel that handles foreign currencies" Or "You need to speak to the international student adviser to make sure that your visa is legitimate and that it will allow you to enroll for courses."

3. PRETESTING

Given that a good back-translation suggests but does not guarantee a good translated version, other methodological steps should be added. One is pretesting. The translated version should be administered to approximately 30 monolinguals, who then answer the questions, complete the standardized test, or go through the experimental procedures, depending on the nature of the research instrument. These pretest participants should be similar to the eventual respondents in the proposed final phases of the research. Pretest participants should be similar in terms of age, gender distribution, amount of formal education, status within their communities, and so forth. With 30 participants, internal consistency reliability estimates can be obtained for questionnaires and tests. Participants can be observed going through the necessary steps in an experiment and adjustments can be made where they have difficulties. With the help of research assistants (perhaps the translators, who should be integrated as members of the research team), participants can be asked what wordings were clear and unclear, what procedures they found difficult, and so forth.

4. EMICS AND ETICS

In any large-scale research project, some original language material will not translate well. This is to be expected, and in many ways researchers should look forward to the challenge of interpreting difficult-to-translate material. As research efforts become more global, the cross-cultural aspects of translation have become even more critically important in developing research of the same, multilinguistic nature. Cross-cultural researchers have laid out important groundwork in demonstrating both potential pitfalls and proposing guidance for developing research across country borders. The idea of universal concepts (etic) or culture-specific concepts (emic) holds particular importance in translation methodology. Any cross-cultural or international study that does not consider the etic/emic issue when developing research instruments or conducting research will create studies of limited validity and value.

Etic approaches to research are studies conducted cross-culturally in order to extract common elements. Hofstede's well-known IBM study is a prime example of an etic approach to cross-cultural research. Etic research has the potential for reaching much broader, and culturally sweeping, conclusions that are often beyond the scope of only one cultural landscape. Etic aspects of language represent ideas that are more universal, with semantics that are more readily translated between cultures.

Emic approaches to research are studies conducted within a cultural system based on data representative of only that cultural landscape. Emic (e.g., unique and regionally specific) constructs of language can create roadblocks to both the interpretation of research results and how clearly the research parameters and ideas are interpreted by the subjects. One reason for difficulties in translation is that a culture's emics, or specific manifestations of complex concepts, have not been clearly identified.

For example, the concept of friendship is universal (an etic), but there are specific aspects of friendship that are emic. In the United States, the term friend is used very loosely and includes recently met acquaintances. In China, the term is used to refer to a smaller number of people who have long-term mutual obligations to each other. The need to differentiate between "older brother" and "younger brother" in Japanese is another example of an emic. Another example can be found in the neo-Melanesian language, which has no word for "desert" because rainfall is plentiful in Melanesia and plants grow in most places. However, there are worn-away lands that are referred to as "nothing places" or "ples nating." In translations of the New Testament, John baptizes in such places rather than in the desert. Clear failures in translation and data

collection can often involve a misunderstanding of another culture's emics, which is a possibility in most phases of cross-cultural research. For this reason, emic manifestations of concepts should be documented and reported. In cross-cultural studies, the reporting often takes the form of what aspects of a concept are shared across cultures and what concepts are culture-specific. Researchers should always remember that it is the combination of etics and emics that advances our understanding of concepts among members of our own and other cultures.

One way of recognizing emic differences, and therefore enhancing the potential of the research in achieving translation and conceptual equivalence, is through the use of back-translation. Translation techniques generally and back-translation techniques in particular have assumed even more importance as researchers attempt to partial out those concepts that are etic in nature from those concepts that are emic.

5. DECENTERING

When cross-cultural researchers are sensitive to the emic/etic distinction, they are careful not to impose their own emics on other cultures. Many concepts in psychology (e.g., personality, intelligence, and worker motivation) reflect the culture in which they were developed and in which research on them took place. When one of the authors gave a lecture on personality to a group of Chinese scholars, they later commented that they found many of the specific terms (achievement orientation, need for power, openness to experience, etc.) difficult to understand since there were few readily available terms in Chinese. They mentioned that they have a much richer language to describe people's relations with others, which would follow from China's status as a collectivist culture. The key point is that personality terms in English and collectivist terms in Chinese may reflect the two culture's emics. To expect an easy translation of them could be a mistake.

When researchers use decentering, they make attempts to find concepts that are readily expressible in the languages under study. No original language version becomes the "center" that has to be translated into less familiar and sometimes clumsy target language terms. Rather, the original language version is open to change as each translated/back-translated version is reviewed. With each review, researchers reword the original until a smooth target language version (as reported by translators or independent bilinguals) and back-translation version are developed. In the case of published standardized tests, this procedure demands that researchers be willing to revise the original wording just as if they were formulating an instrument for the first time. The assumption is that more meaningful information will be obtained from a decentered instrument than from a forced translation of English that has problematic equivalents in the target language.

6. WRITING TRANSLATABLE ENGLISH

If researchers prepare their own materials "from scratch," there are guidelines for writing English that translators may find relatively clear and understandable. These suggestions were developed by Brislin in 1970 as part of a research project on improving translation for cross-cultural research projects. Working with 92 translators, he found that they had a relatively easier time with their task if the original English was prepared according to the following suggestions:

1. Use short sentences of less than 16 words. With longer sentences, translators have difficulties disentangling what subjects go with what verbs.
2. Employ the active voice. Some languages do not have passive voice constructions, and the active voice often allows translators to more clearly identify subjects, verbs, and objects.
3. Repeat nouns instead of using pronouns. Some languages have more pronouns than there are in English, and so the translator may be forced to make choices from alternatives that the researcher does not know about.
4. Avoid metaphors and colloquialisms. Such phrases, part of the socialization of children in a specific culture or geographical areas within a large country, often do not have equivalents in other languages.
5. Avoid the subjunctive (i.e., phrases with "could," "would," and "should"). Some languages do not have easily available terms for the English subjunctive, and so the translator is forced to make approximations.
6. Add sentences to provide context. Expanding on the previous discussion of "friend," there might be a sentence indicating how long people have known each other. This suggestion sometimes leads to longer items, but the items become multisentence items given suggestion number 1.

7. Avoid adverbs telling "where" (e.g., beyond) or "when" (e.g., frequently). These terms often have inadequate direct equivalents and they force translators to make their own decisions on questions such as "how many times is necessary for a translation of a phrase such as 'meeting frequently with my boss?'"

8. Avoid possessive forms when possible. English speakers are familiar with ownership and have extensive experience with its description. This is not the case in cultures in which private ownership of property and goods is not as extensive and in which collectives rather than individuals control ownership.

9. Use specific rather than general terms (e.g., "wheat" rather than "edible plants"). Since cultures categorize specific terms differently, there is no guarantee that a general term will have the same connotation in other languages.

10. Avoid words indicating vagueness regarding some event or thing (e.g., "probably," "maybe," and "perhaps"). They do not always have translation equivalents, and they force translators to make difficult choices among imperfect alternatives in the target language.

11. Use wording familiar to translators. If translators are familiar with English terms, they can indicate the presence or absence of equivalent terms in the target language. This suggestion reinforces the previous recommendation that translators be integrated into the research team.

12. Avoid sentences with two verbs if the verbs suggest different actions. Translators often have difficulty linking the appropriate subject with the appropriate verb.

7. RECENT ENHANCEMENTS OF TRANSLATION METHODOLOGY IN RESEARCH

During the past 15 years, there has been a significant effort to further develop and enhance translation effectiveness, and in doing so this effort has recognized that there are more elements to effective translation than the multistep process described previously. Most of this research and theory development has focused on four basic areas: the initial development of the questionnaire, the style of translation, the use of bilingual or monolingual translators, and the use of quantitative measures to establish item equivalence. A secondary area of theory development concerns the effectiveness of this methodology in qualitative research.

7.1. Questionnaire Development

Research has shown that the literal translation of a data-collection instrument from one culture to another can put in doubt the reliability of the instrument and the study. Of even greater concern is the question regarding the ethics of doing research, and drawing conclusions, from the focus of the dominant culture and not from the focus of the culture being study. Decentering, as described previously, is one approach used to minimize this problem. However, a dual-focus approach to translation has also been offered as a possible solution to these concerns.

The dual-focus method uses a two-pronged approach for developing research instruments. Initially, and most important, a horizontal collaboration of researchers from the indigenous cultures that are being studied is established to develop the original research instrument. After the initial instrument is developed, researchers focus on capturing the concepts of the phenomena being measured in the instruments rather than simply focusing on literal translations. It has been offered that the dual-focus approach to instrument and measure development is superior to other translation methods because the approach is less time-consuming and overcomes many of the concerns regarding the validity of the decentering process. The dual-focus approach may also be better suited for multilingual studies compared to the decentering approach, which has been described as being extremely complex and cumbersome in these instances.

7.2. Translation Style

Although most current approaches to translation have shifted from literal to conceptual translation of meaning, there is a school of thought that advocates that the nature of the translation is in part due to the nature of research being undertaken. Translation can be either asymmetrical or symmetrical. Asymmetrical translations of instruments seek to identify differences between cultures and therefore must remain true to the original language in which the instrument was designed. A literal translation is required, which in turn can create a translated version that may appear unnatural and unusual to the target population being examined. On the other hand, symmetrical translations of instruments seek to identify similarities across cultures. Therefore, the instruments must be recognizable and colloquial in content, with items equally understandable to all the cultures within which the research is being conducted.

7.3. The Use of Monolingual and Bilingual Translators and Subjects

Brislin found that in judging the equivalence of translated material, monolingual or bilingual judges would be equally effective in the discovery of errors of meaning. However, further research has provided additional insight as to what to consider when using translators.

Translators who are familiar with semantic differences are a necessary requirement for cross-cultural studies that collect data from subjects in different regions with different dialects. Cultural perceptions that are held by the translator are also a significant aspect to the translation process, especially when the goal in translation is symmetrical rather than asymmetrical. It is therefore considered optimal to have the team of bilingual translators acquire skills in both languages at different times and in different locations. Expertise in the content of the instrument is also an important attribute to equivalence in translation. Bilingual translators who are familiar with the content of the research and how the instrument will be used develop more effective translations.

Perhaps the most extreme approach to issues of translation and conceptual equivalence is to recruit bilinguals to participate in the research study. Depending on the nature of the research and the type of study population to be examined, using bilingual subjects can enhance data collection quality, especially when the focus of the research is conceptual and comparative.

7.4. Quantitative Measures to Establish Equivalence

The use of translation and back-translation, as well as decentering techniques, is designed to enhance the equivalence of translated materials with original versions. Although effective, the use of decentering techniques becomes much more difficult as different language versions are developed. In response to this problem, certain quantitative techniques can be used after translations are made to further assess the accuracy of these translations. Item response theory provides a method of comparing responses from different translated versions of questionnaires to determine if there is equivalence. This quantitative approach is most effective using dichotomous data. When using polychotomous formats, simultaneous factor analysis can be used to determine equivalence. Pearson product moment correlation, median correlation coefficients, as well as mean scores and standard deviations are also quantitative tools that can be used to determine equivalence between source language questionnaires and target language translations.

8. TRANSLATION AND QUALITATIVE STUDIES

Qualitative research literature has not extensively addressed the significance of equivalent translations given the nature of the research methodology. It is assumed that the face-to-face nature of the communication allows for in-depth probing and clarification of concepts and data between researcher and informant. For this reason, little emphasis has been placed on translation as a critical component of research methodology. However, translation equivalence is a critical concern when international qualitative data are being reduced into specific categories or themes or when subtle and complex phenomena are being studied.

When categorizing data, translation equivalence errors may not necessarily occur in the assignment of translated material into major categories and themes. Rather, these translation errors can frequently appear when the major categories are reduced into more refined and sensitive subcategories and subthemes. It is here that translation equivalence has a significant bearing upon the quality of the research.

Translating complex in-depth interviews from one language to another can create inaccuracies due to lack of word equivalence as well as difficulty in identifying interview filler so as to establish accurate dross rates. Attempting to replicate grammatical style between two grammatically different languages is not only a problem in quantitative research but also problematic in qualitative research. Difficulties increase in trying to understand not only what was said but also how one was saying it, especially when what is communicated is both complex and subtle.

Because interview data can be described in much more complex detail than a question in a questionnaire, the use of multiple, independent translators and back-translation may not be an effective approach for achieving translation equivalence. It has been recommended that when translating qualitative data a single, expert translator should be used. This will allow for consistencies in interpretations of language usage and nuances.

Aside from the translation of completed interviews from one language to another, an additional concern is the use of bilinguals in conducting research interviews. Although bilinguals may have the ability to probe and simultaneously translate for common meaning, there is concern that bilinguals will inconsistently draw upon both languages when conducting these interviews. English words or phrases, for example, are often interjected into foreign language conversation. The multiple meanings that these words or phrases may have to a bilingual interviewer compared to a monolingual interviewer may also put into doubt the validity of the data and data analysis.

Like quantitative research, qualitative research can be impacted adversely by lack of equivalence in translation. Although these problems may impact different aspects of the research, they still need to be considered and addressed both in the planning and in the execution of the research study. It is also important to be aware that because problems of translation may impact different aspects of qualitative research, the use of standard translation methodologies in quantitative studies may not be appropriate.

9. NEW CHALLENGES IN TRANSLATION: THE PURSUIT OF EQUIVALENCY IN HIGHER ORDER ISSUES

During the past two decades, the need for better and higher quality translations of questionnaires and other data-collection instruments has become increasingly important to cross-cultural and international researchers. As more attention is focused on the translation methodology and some issues of conceptual equivalency, new problems occur. These problems focus on the more subtle, higher order aspects of translation and conceptual equivalency.

9.1. Nonverbal and Other Special Translation Issues

The importance of nonverbal cues in the overall communication process is well documented. Some estimates indicate that more than 80% of receiver information is acquired through nonverbal cues. Pitch, tempo, loudness, hesitation, facial cues, hands and hand gestures, timing, silence, symbols, and even dress can add meaning to a translated communication. These nonverbal cues are even more important when there is significant divergence between the host culture and targeted culture during translation. Research has supported the notion that in-group advantages for cultural members are present for correctly interpreting and decoding nonverbal communication and emotion recognition. This in turn suggests that out-group members or expatriates may encounter additional challenges within these more nuanced areas of translation.

9.2. Professional Jargon, Metaphors, and Idioms

Clearly, in-group/out-group misinterpretations can also occur outside of nonverbal communications. Professional jargon, idioms (a phrase or expression having a different meaning from the literal), and slang are all types of communication that require careful consideration and more than just a literal translation in order to achieve equivalence in meaning between two languages. Strategies that can be used to consider these special cases include finding ways to neutralize the text; attempting to find an exact equivalent of the original jargon, metaphor, or idiom; selecting other jargon, idioms, or metaphors that would express a similar sense; or replacing the jargon, idiom, or metaphor with an approximate literal phrase.

10. FUTURE DIRECTIONS OF TRANSLATION AND TRANSLATION MODELS

The 21st century will present new challenges to achieve effective translations as the world becomes smaller and more reliant on technology. The opportunities for faster and more insightful translation of data instruments and research data will also create new problems and challenges for researchers, particularly as researchers shift their attention from identifying universal etics to recognizing culturally specific emics.

10.1. The Impact of Globalization

Language can be viewed as one of the more important gateways to culture that has challenged human interaction since our earliest history, as exemplified through accounts such as the *Tower of Babel*. As an integral and prominent artifact of culture, language, and the process of understanding known as semantics,

plays a critical role in a world community that has been moving toward globalization. Recent trends such as the globalization of business and trade, worldwide communication networks, new technology, improved transportation, and geopolitical economic unions will continue to increase the incident of foreign language contact and thus the need for effective, efficient translation. These needs also drive academia toward new studies of language, semantics, and translation.

10.2. Technology and Translation

In developing a forward-looking view of translation, any treatment would be remiss by not discussing the use of technology. New technology, including fuzzy logic, neural networks, and parallel processing (i.e., artificial intelligence), is developing in ways that will radically alter the reality of translation as we know it today. The task of holistic, global translation, taken from an individual's perspective, is indeed daunting. There are 3000–6000 languages globally (depending on the source cited), with approximately 200 of these being spoken by 1 million people or more and approximately 100 serving as official, national languages. Given that even the best human translators are able to work effectively with only a handful of these languages, a significant gap remains. Translation technology not only promises to functionally assist the facilitation of communication between individuals, groups, and organizations but also may help to eventually preserve more fringe languages. By acting as an "intellectual crutch," translation technology may discourage the learning of a new language, entrenching the native tongue.

Currently, textual translation programs and software are integrated and used on the Web in business and in other organizational contexts. For example, the European Union uses translation software to handle approximately 10% of its translation duties. Large, global organizations are predicting a significant increase in the total amount of material that will need to be localized to fit regional language patterns—as much as sixfold by 2005. These trends, coupled with a stagnating number of human translators, are driving forces within the translation industry, which has been growing in gross revenue by 15–20% per year. Nevertheless, we are just beginning to use technology as an effective tool for translation and as an effective leverage for managing language.

Automated translation software has been used since the Cold War era, with some of the more modern, advanced programs able to translate in 1 hour, with 90% accuracy, what might take a human 2 or 3 days. For example, synchronous automated translation systems (SATS) represent highly advanced technology that will allow seamless translation without any perceptible time delay. SATS-based systems are also expected to handle basic aspects of idiom and will be able to alert communicators of potentially vague language that would have a high probability of multiple meaning, thus prompting for more clarification. Advanced SATS systems are projected to be mainstream and in the market by 2015. One of the main trials for this type of advanced translation system will be how effectively it can handle a variety of unique translation issues, such as interaction patterns between low- and high-context cultural societies and the relative importance of contextuality in the communication process.

11. CONCLUSION

During the past 30 years, the need for effective and more precise translations of research materials has become a recognized requirement for high-quality international or cross-cultural research. The challenge of achieving conceptual equivalence and meaningful understanding among peoples of different cultural backgrounds and language has taken the field of translation far beyond "finding the same word in a different language." This article discussed the problems that researchers have identified in achieving quality translation of research materials and the approaches needed to achieve the level of desired quality. The translation/back-translation, decentering model first proposed in the 1970s provided the foundation for which virtually all future approaches and concerns regarding quality translations have been addressed. Once conceptual equivalence in language is achieved, new challenges arise regarding how to achieve this conceptual equivalence with multiple dialects and contexts.

During the past 10 years, there have been many modifications and enhancements of the translation model as it has become more widely used in fields outside of psychology. Nursing, public health, leadership, and international business studies have all recognized the general power of the model as well as identified ways, and circumstances in which, it can be further enhanced. This article also identified some of the new frontiers and future challenges awaiting researchers as they attempt to achieve greater conceptual equivalence that considers not just verbal language

translations but also nonverbal, semantic, and contextual equivalence. Finally, the article discussed the future of translation with the use of technology and what that may mean to the field.

The goal of all these efforts is simple: to enable people to better understand one another. In doing so, we will be better able to understand why people in other countries and cultures do the things that they do, and in the process we will be able to understand why we behave the way we do. Ultimately, the true purpose of effective equivalent and conceptual translation is to help us understand and, in doing so, bring us closer together.

See Also the Following Articles

Conformity across Cultures ■ Cross-Cultural Psychology, Overview ■ Literacy, Improvement of

Further Reading

Babcock, R., & Du-Babcock, B. (2001). Language based communication zones in international business communication. *Journal of Business Communication*, *38*(4), 372–412.

Berry, J. (1969). On cross-cultural comparability. *International Journal of Psychology*, *4*, 119–128.

Bontempo, R. (1993). Translation fidelity of psychological scales: An item response theory analysis of an individualism–collectivism scale. *Journal of Cross-Cultural Psychology*, *24*(2), 149–166.

Brislin, R. (1970). Back-translation for cross-cultural research. *Journal of Cross-Cultural Psychology*, *1*, 185–216.

Brislin, R. (1986). The wording and translation of research instruments. In W. Lonner, & J. Berry (Eds.), *Field methods in cross-cultural research* (pp. 137–164). Beverly Hills, CA: Sage.

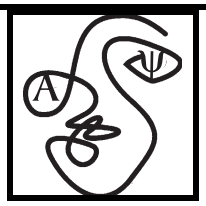
D'Andrade, R. (2002). Cultural Darwinism and language. *American Anthropologist*, *104*(1), 223–232.

Erkut, S., Alarcon, O., Coll, C. G., Tropp, L., & Garcia, H. A. V. (1999). The dual-focus approach to creating bilingual measures. *Journal of Cross-Cultural Psychology*, *30*(2), 206–218.

Freivalds, J. (1999). The technology of translation. *Management Review*, *88*(2), 48–52.

Lehman-Wilzig, S. (2001). Babbling our way to a new Babel: Erasing the language barriers. *The Futurist*, *35*(3), 16–23.

Serebryakova-Collins, I. (1998). Here's a twist: Talk about translation. *Business Communication Quarterly*, *61*(4), 109–111.



Transportation Systems, Overview

Barry H. Kantowitz and John M. Sullivan

University of Michigan Transportation Research Institute, Ann Arbor, Michigan, USA

1. Introduction
2. Operator Workload and Distraction
3. Driver Vision
   Further Reading

GLOSSARY

central fovea A central region of the retina extending about 2 degrees in diameter of visual angle and exclusively composed of densely packed cone receptors. It is the part of the retina that supports seeing fine details and color.

driver distraction A reduction in a driver's attention that occurs when some secondary task interferes with the primary task of maintaining safe control of a vehicle in motion.

human factors/human factors engineering The discipline that tries to optimize the relationship between people and technology.

mesopic sensitivity A mixture of scotopic and photopic vision at low light levels, between approximately 0.01 cd/m^2 and 3 cd/m^2, a common level found in nighttime driving.

National Highway Traffic Safety Administration (NHTSA) A unit of the U.S. Department of Transportation that is responsible for reducing deaths, injuries, and economic losses resulting from motor vehicle crashes.

photopic sensitivity Chromatic, cone-mediated vision used at high light levels (greater than about 3 cd/m^2).

pilot workload An intervening variable that modulates the tuning between the demands of the environment and the capacity of the operator.

scotopic sensitivity Monochromatic, exclusively rod-mediated vision occurring at very low light levels (less than about 0.01 cd/m^2).

underload A potential problem in highly automated modern aircraft in which most routine operations are carried out by the plane itself, which could let the flight crew become distracted or inattentive.

United States Department of Transportation (USDOT) Part of the executive branch of the U.S. government responsible for ensuring fast, safe, efficient, accessible, and convenient transportation systems for the American people.

Driving is a dangerous activity that is often taken for granted without understanding the risks of being on the road in a vehicle. Psychology can improve transportation systems in two ways. First, it provides explanations to help the transportation professional understand operator behavior. Second, its models can lead to effective countermeasures to decrease risk. These points are illustrated with discussions of operator workload and distraction and driver vision.

1. INTRODUCTION

Driving is a dangerous activity that can be hazardous to one's health; it can even kill. The World Health Organization estimates that road crashes are the third most important worldwide threat to public health. Indeed, the April 2004 World Health Day featured automobile crashes as its major topic. While most fatalities occur in developing countries, the United States also has a deplorable safety record, with over 42,000 fatalities in

©2004 Elsevier Inc.
All rights reserved.

2003. Someone dies in a car crash about every 12 minutes on average. A person is injured in a car crash about every 10 seconds. Automobile crashes are the most likely cause of death for young adults in the United States.

It is difficult to understand why society tolerates this high cost of mobility. If the aviation industry had a similar safety record, no airplane would ever leave the ground. Even considering the terrorist hijacking and crashing of four airplanes on September 11, 2001, flying is still much safer than driving. The risk of a fatality driving only 18 kilometers on a rural interstate highway equals the risk of any non-stop flight, regardless of distance, on a major airline.

Fatal crashes are a very low-frequency event for any single individual; no one reading this text has died in a fatal car crash. Therefore, many people believe that a car crash is an event that involves someone else. This comfortable belief disappears when a friend or a loved one is killed or injured in a car crash.

Most car crashes are due to human mistakes rather than failure of equipment. It is difficult to measure precisely what percentage of crashes is due to driver error because the basic data come from police reports after the crash. A dead or unconscious driver cannot explain the circumstances of a crash, and a live driver may not readily admit making an error. A recent University of Michigan Transportation Research Institute (UMTRI) analysis of a fatal accident database maintained by National Highway Traffic Safety Administration (NHTSA) estimated that approximately 60% of fatal crashes have human error as a key factor. This estimate is conservative; other estimates have ranged as high as 70 to 90%. Drivers engage in all manner of risky behaviors, such as drinking alcohol, not using seatbelts, and talking on cell phones. All too often, such risk is translated into death and injury on the highway.

At least part of the crash risk is a consequence of technological progress. Prior to the 20th century, surface transportation was largely accomplished with the aid of horsepower. This placed strict limits on the speed and weight of a conveyance and the amount of time the conveyance was in service. Moreover, from time to time, a horse would display a degree of intelligence that ensured its rider returned home whether or not the rider was fully conscious. Horsemanship was also considered a skill developed over time—it was decidedly dangerous for an unpracticed rider to mount certain horses.

The modern automobile has removed many of the limitations on horse-powered travel. Automobiles can achieve higher speeds, carry more cargo, and operate more reliably and more consistently than horses. More automobiles are on the road than were horses, and drivers of widely different levels of skill and education are licensed to operate them.

These technological advances have indirectly raised the performance demands on automobile drivers. At high speeds, the available judgment time to take an evasive action, for example, is shorter, as is the time to correct an error. With heavier vehicles, momentum (kinetic energy) is higher, so collisions are more destructive. With more roadway users of varying levels of skill, there is more potential for interaction and more complicated decisions for the driver to make. With the aid of automotive lighting and fixed roadway lighting, travel at night is now both feasible and commonplace but also raises new questions regarding human performance in varied environments.

While technology enhances the risk of human death and injury by multiplying forces well beyond that produced by an unaided human, it also offers the potential to mitigate such damage. For example, in the United States, an Intelligent Vehicle Initiative research program attempts to use new technology to improve road safety by improving the understanding of driver distraction in order to develop new methods to measure the effects of in-vehicle technology on driver performance, to create new human factors guidelines for equipment design, and to develop new integrated approaches to reducing driver distraction caused by in-vehicle information systems.

Psychology, as the discipline that purports to explain and understand human behavior, has much to contribute to improving transportation systems. Its contributions are made in two ways. First, psychology provides explanations that help transportation professionals understand driver behavior. People are not mathematically rational and often satisfice rather than optimize. Yet their behavior is consistent and can be predicted. The following sections emphasize understanding the vehicle operator. Second, psychology can use its understanding to help construct countermeasures that will improve transportation safety. While many talented and dedicated transportation engineers have made substantial contributions to traffic safety based upon their experience and observations, these professionals are seldom well-trained in the subject matter and methods of psychology. Psychology, with its focus upon human behavior, can provide additional insights for developing countermeasures. The following sections offer psychological perspectives on effective countermeasures.

The remainder of this overview article uses two topics, pilot workload/driver distraction and driver

vision, to illustrate briefly how psychology is applied to transportation systems. Both understanding and countermeasures are discussed. Greater detail about specific transportation sub-systems will be found in the articles that follow this overview.

2. OPERATOR WORKLOAD AND DISTRACTION

Regardless of the type of vehicle being controlled, demands on the operator, whether a pilot of an airplane or a driver of a passenger car, must fall within his or her mental and physical limitations. In modern transportation systems, human physical limits are seldom exceeded, but the cognitive load placed on an operator can often be incongruent with human capabilities. This psychological phenomenon has been extensively studied in aviation and is currently a major research topic in driving.

2.1. Pilot Workload

Airplane pilots are highly trained and must demonstrate continued skill levels and sound health on a repetitive basis to maintain their licenses. Simulators are used extensively to practice emergency skills, and a rigorous system of checks and balances ensures the safety of the flying public. This system works well; flying on a major airline is far safer than driving on a highway.

One of the reasons for this high margin of safety is decades of research on pilot workload. An airplane pilot cannot pull over to the side of the road, as can a driver, when some problem develops. Pilot workload is defined as an intervening variable that modulates the tuning between the demands of the environment and the capacity of the operator. It is easy to understand that if workload is excessive, the pilot will be unable to make correct decisions in a timely manner, putting the aircraft at risk. Modern airplanes have extensive automation to reduce pilot workload. But if workload is too low, the pilot may be "out of the loop," perhaps even asleep, and cannot respond correctly in a timely manner. Underload can be a problem in highly automated modern aircraft, in which most routine operations are carried out by the plane itself. Pilots sometimes joke that the aircrew of the future will be one human pilot and a dog; the dog's job will be to keep the pilot from touching the controls. Some pilots believe that their manual flying skills are being degraded by the frequent use of automated systems. Indeed, a 1988 review of pilot workload by Kantowitz and Casper concluded, "It would be tragic and ironic if our current efforts to measure pilot workload succeeded, only to be faced with a new generation of aircraft where pilot workload was so low that nobody bothered to measure it at all" (p. 184).

Fortunately, the theoretical understanding of pilot workload derived from cognitive models has been applied to the issue of automation. Automation can be implemented in a sensible way that takes pilot capabilities into account, or in a clumsy way that makes the pilot's tasks more difficult to accomplish. The pilot should always be able to understand what the automation is doing at any moment. Some cockpit designs make this easier than others. For example, in Boeing aircraft, the throttle moves in response to automation commands, giving the pilot useful visual feedback. However, in Airbus designs, to save weight the throttle does not move, making it more difficult for the pilot to discover what the automation is doing.

Automation tends to fail in a sharp, sudden manner compared to the graceful degradation of human pilots. When automation can no longer cope with a situation, it turns off, leaving the pilot to deal with the problem. An incident that illustrates this occurred when a pilot made an initial error in transferring fuel from the wing tanks to a center tank. If the weight of fuel is not evenly distributed, the plane can roll over. The automation did correct for this error up to a point, but eventually the imbalance became too great. At that point, the automation turned off, and the pilot suddenly realized there was a significant problem. Had the plane been under manual control, this problem would have been detected much earlier by the crew. Although automation has improved aviation safety, it can produce new problems if the needs of the human pilots are not considered most carefully.

2.2. Driver Distraction

Many of the theoretical and methodological tools used to learn about pilot workload can be, and are being, applied to the problem of driver distraction. Driver distraction occurs when a secondary task, such as tuning a radio, eating a pizza, entering data for an in-vehicle navigation system, or talking on a cell phone, interferes with the primary task of maintaining safe control of a vehicle in motion.

Driver workload and distraction are often studied by comparing some measure of driving performance with and without secondary tasks. It is important to remember that workload and distraction are theoretical

constructs that cannot be directly observed. Instead, they must be inferred from changes in performance, which can be observed and measured directly. Thus, there must be some experimental conditions that contrast single-task (only driving) and dual-task (driving plus other tasks) performance before inferences about workload and distraction can be generated. The typical outcome of such experiments is that adding most secondary tasks detracts from driving performance. The practical conclusion from such results is that effective countermeasures will limit the driver's ability to perform extraneous tasks, especially when the workload imposed by driving is high. Driving workload can be increased, for example, when sharp radius curves are negotiated, when the driver is fatigued, or when traffic density is substantial.

A very clever field study was able to relate secondary-task reaction time to road safety. In 1986, Harms showed that driver workload on 100-meter segments of roadway covaried with the number of reported accidents on these segments. Decreased performance on a secondary task, given constant primary-task performance, is a common index of driver workload. While many studies have demonstrated this in driving simulators and on test tracks, Harms was one of the first researchers to link workload and safety on actual highways.

Talking on a cell phone is a prime example of a secondary task that reduces the driver's ability to control a vehicle safely. Some states have banned the use of hand-held phones on the road, but a hands-free phone is also quite dangerous. An epidemiological study found that the relative risk of injury and collisions is 38% higher for cell phone users. Studies of vehicles on test tracks have also demonstrated the distraction effects of cell phone usage. Drivers who need to use a cell phone in the car should pull over to a safe place (a parking lot rather than the side of an interstate highway) to talk.

It is certainly possible to use cognitive models of the driver to invoke new technology that will build adaptive interfaces to limit secondary tasks, especially telematic secondary tasks from devices inside the vehicle. For example, in-vehicle navigation systems know when the driver is approaching a dangerous curve, and could then turn off the driver's cell phone. It seems likely that these new interfaces, sometimes called workload managers, will appear soon in new vehicles as effective countermeasures for excessive in-vehicle information loads. A recent on-road test of such a device used sensors that monitored the roadway and vehicle dynamics to dynamically estimate driver workload. When workload was too high, cell phone calls were automatically diverted to a telephone mailbox without asking the driver's permission. Satisfaction with this demonstration system was high, suggesting that commercial systems can be developed to improve driver safety and comfort.

Many of the new in-vehicle information systems being deployed in ground vehicles have been developed not by large automotive companies, but by smaller suppliers, who often lack substantial in-house knowledge and capability in the fields of psychology and human factors. Fortunately, researchers have responded to this need by developing human factors guidelines funded by the United States and European governments that are available to all designers. For example, one highly used set of guidelines for advanced traveler information systems contains 75 specific design rules that can be applied by users who lack formal training in human factors. (These and other guidelines can be found on the Web by first searching for Turner Fairbanks Highway Research Center and then for human factors guidelines.)

3. DRIVER VISION

Key ingredients for safe driving include staying on the road, maintaining safe distances to other vehicles, recognizing and avoiding obstacles in the roadway, predicting other vehicle trajectories, reading and understanding traffic signs, and obeying traffic laws. Vision plays a critical role in providing input to the earliest stages of the complex chain of decisions and actions that influence driving behavior.

There are many important links between what a driver sees and what a driver does. A driver sees lane markings in the roadway, and steers to the center of the lane; sees a stopping vehicle ahead, and judges closing speed; sees a child at the side of the road, and slows down; sees a gap open up, and merges into traffic; sees a stop sign, and applies the brake; sees a detour sign, and plans an alternate route; sees an exit sign, and merges right in preparation for an exit. Seeing plays such an important role in driving that the desire to quantify it is difficult to resist. In a 1996 paper, Sivak noted that while it is often suggested that 90% of driving is visual, no real quantitative basis actually exists for this number. What can be quantified is the change in crash risk when visibility is reduced. In a 1996 article, Owens and Sivak found that reduced visibility in darkness doubles and quadruples the risk of some types of fatal crashes compared to the same crashes in daylight.

3.1. Basic Visibility

In discussing driving and vision, a distinction should be made between sensation and perception. To see an object, light receptors in an observer's eye must generate sufficient neural energy to signal the object's presence. This initial transduction of light energy to neural activity is principally the domain of sensation. However, even when adequate sensory registration of an object occurs, whether an object is actually seen and recognized also depends on the context of the visual scene and the knowledge, readiness, and experience of the observer. For example, objects are more quickly and reliably detected in an uncluttered scene than in a cluttered one; expected objects are more quickly and reliably detected than unexpected ones. Perception concerns what a person finally sees after registering sensory input.

Considering sensation first, if the receptors in the eye are insensitive to a stimulus, there is nothing for the perceptual system to interpret. The light sensitivity of the eye is dynamic and is influenced by the response characteristics and distributions of two classes of photoreceptors in the retina—rods and cones. Rods are more sensitive than cones and enable vision under low light levels, but they are disabled under bright conditions. Rods are absent in the central fovea of the retina, and rod-mediated vision is monochromatic. Cones are less sensitive to light than rods and are active under bright viewing conditions. They are densest in the central fovea and decline in number toward the periphery; there are three classes of cones with differing sensitivities to wavelengths of light, enabling color vision. To partially distinguish these differences in the response of the eye, vision has been identified as scotopic (monochromatic, rod-mediated vision) at low light levels (less than about 0.1 cd/m^2), photopic (chromatic, cone-mediated vision) at high light levels (greater than about 3 cd/m^2), and mesopic (a mixture of scotopic and photopic vision) at light levels in between. Light levels for nighttime driving conditions are generally in the mesopic range, although many objects illuminated by headlamps are in the photopic range, and unlit objects are in the scotopic range.

The implications for night driving are that objects that are looked at directly (i.e., exclusively with cones in the central fovea) require at least 3 cd/m^2 to be seen clearly. Consequently, most objects require supplemental illumination to be seen with the same clarity as in daylight. One solution would be to flood the roadway with as much light as possible. However, such a strategy is both impractical and potentially disruptive to other road users, who may be temporarily blinded or discomforted by the light. Low-beam headlamp designs address this problem by distributing light about the roadway in a way that attempts to balance the vehicle operator's illumination needs with the need to minimize glare to other road users. Light is aimed down, below the horizon, and to the right side of the road. This limits forward seeing distance and, particularly at high speed, leaves little time to respond to an obstacle in the roadway. A second, switched, high-beam system in intended to address this limitation by sending more light down the roadway. In theory, the driver switches between the two beam systems as the driving conditions require. In practice, it was found that only about less than half of the drivers in a Michigan study used their high-beam headlamps when they should.

Perhaps this happens because low-beam headlamps appear to provide sufficient illumination on modern roadways, which are marked by reflectorized lane striping and other supporting illumination. Drivers may not be aware that there is a visibility problem at all, judging their visual competence by their ability to maintain lane position. Unfortunately, while a driver's steering and lane-keeping abilities appear to suffer little under low-beam illumination, detection of unreflectorized, low-contrast objects appears to be selectively degraded. Leibowitz, Owens, and Tyrrell have suggested that this situation leads drivers to routinely overdrive their headlamps. That is, their effective stopping distance exceeds their seeing distance. Consistent with this view is the fact that drivers involved in fatal nighttime pedestrian crashes often report not seeing the pedestrians at all.

A variety of countermeasures have been devised to extend a driver's basic visual sensory capabilities while holding the problem of glare at bay. such countermeasures include night-vision systems that render the invisible infrared radiation reflected or emitted by a roadway object as a visible image on a display, methods of adaptively redistributing the light from headlamps to optimize visibility in a variety of driving conditions, and more widespread use of reflectorized clothing. At the same time, countermeasures to address the problem of glare appear to dominate the discussion of roadway visibility, perhaps because there is more public awareness of glare than of poor visibility.

3.2. Conspicuity

Even when sufficient light is available to make a target object visible, it may be overlooked. A target's

detectability is also strongly influenced by visual context. A target is detected more slowly in complex visual scenes than in uncluttered visual scenes. Visual attributes that are unique to a target object can make it more conspicuous. Pedestrians who wear reflectors on moving limbs can be recognized at greater distances in the dark than when the same reflectors are worn on the torso. Limb motion heightens conspicuity. The advantage disappears, however, if the pedestrian is embedded in a cluttered visual environment.

Driver expectation also interferes with detection. Pedestrians are detected at greater distances in conditions in which they are expected (e.g., in crosswalks and urban intersections), and at shorter distances when they are not (e.g., along rural limited-access highways at night).

Various measures have been used to heighten the conspicuity of some roadway objects in specific situations. Frequent daytime collisions involving motorcycles have led to regulations requiring use of daytime running lamps on motorcycles. Fatal collisions with tractor trailers have prompted regulations requiring use of reflectorized sheeting materials or lights to mark the rear and side areas of trailers. The problem of rear-end collisions has been addressed with the use of the center high-mounted stop lamps to more clearly portray the action of a braking vehicle. And pedestrian visibility has been enhanced with the use of reflectorized materials in shoes, backpacks, and clothing.

See Also the Following Articles

Accidents in Transportation ■ Aviation ■ Driving Safety ■ Models for Transportation ■ Traffic Safety Assessment ■ Travel Behavior and the Environment

Further Reading

Birch, S. (2001). Adaptive front lighting. *Automotive Engineering*, *109*, 39–42.

Campbell, J. L., Carney, C., & Kantowitz, B. H. (1999). Developing effective human factors design guidelines: A case study. *Transportation Human Factors Journal*, *1*, 207–227.

Gish, K. W., Staplin, L. K., & Perel, M. (1999). Human factors issues related to use of vision enhancement system. *Transportation Research Record*, *1694*, 1–9.

Hancock, P. A., Lesch, M., & Simmons, L. (2003). The distraction effects of phone use during a crucial driving maneuver. *Accident Analysis & Prevention*, *35*, 501–514.

Harms, L. (1986). Drivers' attentional responses to environmental variations: A dual-task real traffic study. In A. G. Gale *et al.* (Eds.), *Vision in vehicles*. Amsterdam: North Holland.

Herslund, M. B., & Jorgensen, N. O. (2003). Looked-but-failed-to-see errors in traffic. *Accident Analysis & Prevention*, *35*, 885–891.

Kantowitz, B. H., *et al.* (2004). Development of critical knowledge gaps and research efforts. FHWA Technical Report, in press.

Kantowitz, B. H. (1988). Defining and measuring pilot mental workload. In J. R. Comstock, Jr. (Ed.), *Mental-state estimation 1987* (pp. 179–188). Hampton, VA: National Aeronautics and Space Administration, Scientific and Technical Information Division.

Kantowitz, B. H., & Campbell, J. L. (1996). Pilot workload and flightdeck automation. In R. Parasuraman, & M. Mouloua (Eds.), *Human performance in automated systems* (pp. 117–136). Mahwah, NJ: Lawrence Erlbaum Associates

Kantowitz, B. H., & Casper, P. A. (1988). Human workload in aviation. In E. Wiener, & D. Nagel (Eds.), *Human factors in aviation*. New York: Academic Press

Kantowitz, B. H., & Simsek, O. (2000). Secondary-task measures of driver workload. In P. Hancock, & P. Desmond (Eds.), *Stress, workload, and fatigue* (pp. 395–408). Mahwah, NJ: Erlbaum.

Laberge-Nadeau, C., Maag, U., Bellavance, F., Lapierre, S. D., Desjardins, D., Messier, S., & Saidi, A. (2003). Wireless telephones and the risk of road crashes. *Accident Analysis & Prevention*, *35*, 649–660.

Langham, M. P., & Moberly, N. J. (2003). Pedestrian conspicuity research: A review. *Ergonomics*, *46*, 345–363.

Leibowitz, H. W., Owens, D. A., & Tyrrell, R. A. (1998). The assured clear distance ahead rule: Implications for nighttime traffic safety and the law. *Accident Analysis and Prevention*, *30*, 93–99.

Luoma, J., Schumann, J., & Traube, E. C. (1996). Effects of retroreflector positioning on nighttime recognition of pedestrians. *Accident Analysis and Prevention*, *28*, 377–383.

Mace, D., Garvey, P., Porter, R. J., Schwab, R. N., & Adrian, W. (2001). *Countermeasures for reducing the effects of headlight glare*. Washington, D.C.: The AAA Foundation for Traffic Safety.

Macmillan, N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide*. New York: Cambridge University Press.

Moberly, N. J., & Langham, M. P. (2002). Pedestrian conspicuity at night: Failure to observe a biological motion advantage in a high-clutter environment. *Applied Cognitive Psychology*, *16*, 477–485.

Morgan, C. (2001). *The effectiveness of retroreflective tape on heavy trailers*. NHTSA Technical Report DOT HS 809 222. Washington, D.C.: National Highway Traffic Safety Administration.

Owens, D. A., & Sivak, M. (1996). Differentiation of visibility and alcohol as contributors to twilight road fatalities. *Human Factors*, *38*, 680–689.

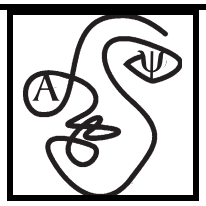
Piechulla, W., Mayser, C., Gehrke, H., & Konig, W. (2003) Reducing drivers' mental workload by means of an adaptive man-machine interface. *Transportation Research Part F*, *6*, 233–248.

Resendes, R., & Martin, K. H. (2003). Intelligent vehicle initiative 2002 annual report-saving lives through advanced vehicle safety technology. FHWA-OP-03–101. Washington, DC: U.S. Federal Highway Administration.

Sivak, M. (1996). The information that drivers use: Is it indeed 90% visual? *Perception*, *25*, 1081–1089.

Sivak, M., & Flannagan, M. J. (2003). Flying and driving after the September 11 attacks. *American Scientist*, *91*, 6–8.

Sullivan, J. M., Adachi, G., Mefford, M. L., & Flannagan, M. J. (2004). High-beam headlamp usage on unlighted rural roadways. *Lighting Research and Technology*, *36*, 59–67.



Travel Behavior and the Environment

Tommy Gärling
Göteborg University, Göteborg, Sweden

1. Introduction
2. Determinants of Travel
3. Transportation Systems
4. Measures for Reducing Private Car Use
Further Reading

GLOSSARY

adaptation of car use Changes such as more efficient car use, suppression of car trips, or a switch to other travel modes.
car-use reduction Decrease in car use by households and individuals.
environmental threats Air pollution, noise, accidents, congestion, and extensive use of land caused by motorized traffic.
intelligent transportation systems Applications of information technology for the substitution of travel or for supporting travel-related decisions.
transportation Movement of people and goods.
travel choice Choices to travel or not, where to travel, how to travel, and when to travel for single trips or multiple trips or tours.
travel demand Demand for travel derived from biological needs, social obligations, and personal desires to engage in activities.
travel demand management Policy measures aimed at reducing or changing demand for travel.

In Western industrialized societies, effective urban transportation systems have evolved. A major factor in this evolution is the versatility of the private car. However, current increasing trends in private car use in urban areas are becoming a serious threat to the environment and people, thus outweighing the benefits. How can private car use be reduced? This article reviews research in environmental and transport psychology that addresses this question.

1. INTRODUCTION

In the interest of achieving effective urban transportation with positive welfare effects, Western industrialized societies have in many ways promoted the development of the private car. Only recently has it become evident that the increasing trends in private car use have serious detrimental effects to the environment and people that outweigh the beneficial effects. How can private car use be reduced? This urgent question is being addressed by research on transportation and the environment. What answers does this research provide? This article reviews what is known from research in environmental and transport psychology about the determinants of private car use, what are likely to be effective means for reducing private car

©2004 Elsevier Inc.
All rights reserved.

use, and what the short- and long-term consequences may be of reducing private car use.

2. DETERMINANTS OF TRAVEL

2.1. Travel Demand

Travel is a derived demand. Fulfilling biological needs, social obligations, and personal desires in human environments requires that people travel. In itself, however, travel is a cost or friction. Only occasionally does travel become a personal desire. In such cases, travel is a goal, not a means for attaining a goal.

The way a society is spatially organized is an essential determinant of demand for travel. The boost in motorized travel is primarily a consequence of two historical trends. One is urban sprawl. Cities have grown in size to accommodate an increasing population, thereby increasing demand for motorized travel. Another trend is that lifestyle changes have increased the complexity of travel demand far beyond the demand for home-to-work journeys on weekdays. Existing alternative travel modes lack the versatility of private cars, which is necessary to satisfy a complex travel demand.

2.2. Travel Choice

Although people's travel options are to a large degree determined by the spatial organization of a society in conjunction with the available transportation means, some degree of freedom always exists in choosing to travel or not, where to travel, how to travel, and when to travel. An extensive research agenda has provided knowledge of how travel choices are made.

Traditionally, research has focused on single-trip choices of travel mode and destination. Travel time and monetary costs are important determinants. Aversion toward wait time, uncertain delays, and changes are other recurrent research findings. An important distinction is between habitual, impulsive, and planned travel. Commuting to work is an example of habitual travel. Shopping trips are more frequently impulsive. Long-distance travel is more meticulously planned than any other travel. Prechoice information search and deliberation have been shown to be reduced when travel becomes habitual. It has also been demonstrated how easily a car-use habit is developed and generalized across situational contexts.

Car travel frequently has multiple purposes. A recent aim of research is to understand interrelated travel choices. The unit of analysis is then a tour or trip chain—for instance, choices related to all trips from home and back during a day. Modeling interrelated travel choices has been an impetus to improve the traditional mathematical–statistical tools. Several generations of computer simulations have therefore been developed. The most recent generation models how people learn about and represent the environment as well as how they learn about options with positive outcomes. In computer simulations, realism of a travel plan replaces the traditional criterion of utility maximization under generalized cost constraints. Given a complex household agenda of activities (work, shopping, and leisure activities, including chauffeuring children), speed of travel is the single most important property of the transportation system. Although speed of car travel has slowed down in many urban areas, it is still the fastest travel mode and therefore likely to be chosen.

Both traditional and recent research has tended to downplay feelings of personal freedom, joy, security, and status as determinants of the choice of the private car. Cost factors may also play a more important role than assumed. Some of the costs of driving are nontransparent and thus may be largely discounted. Conversely, the capital cost of a car is not perceived as a sunk cost that should be discounted. Although the cost of an alternative is lower compared to the variable costs of driving, the additional cost for the alternative is psychologically unnecessary.

3. TRANSPORTATION SYSTEMS

3.1. Properties

Systems for ground transportation in urban areas differ in several respects. The basic means of transportation is walking. Another nonmotorized travel mode is cycling. All faster travel modes are motorized. To the users, speed, frequency of service, monetary costs, and comfort are the most important properties of transportation systems. An additional user-related concept is multimodal or "seamless" transportation. The concept refers to combinations of travel modes that facilitate convenient transportation from door to door. Developing such transportation systems is promising and is likely to enhance the attractiveness of alternatives to the private car.

3.2. Environmental Effects

The worldwide increase in motorized transportation in urban areas is expected to continue. The number of cars is increasing at a faster rate than the human population so that in 2050 more than 3 billion motor vehicles, compared to today's 600 million, are estimated to be in operation.

Conventional gasoline-powered cars emit carbon monoxide (CO), nitrogen oxides (NO_x), sulfur dioxide (SO_2), volatile organic compounds (VOCs), and carbon dioxide (CO_2). Car traffic produces nearly two-thirds of all CO emissions, approximately one-third of CO_2 emissions, more than one-third of NO_2 emissions, and slightly more than one-fourth of emissions of VOCs. Substances are also released from the rubber in tires. These effects are strongly associated with a variety of significant environmental damage and human health problems. Depletion of the ozone layer, acidification, and global warming are examples of the former; respiratory and cardiovascular diseases are examples of the latter.

Other adverse effects of car use include noise, accidents, congestion, and extensive use of land for roadways and parking spaces. It is estimated that approximately 20% of the urban population is exposed to annoying traffic noise. Despite many attempts to improve traffic safety worldwide, fatalities and injuries are increasing. Traffic congestion slows car traffic, resulting in huge additional societal costs of urban transportation. In an attempt to reduce if not solve the congestion problem, increasing use of land for new infrastructure has encroached on agricultural land that is needed to produce food for the two-thirds of the world population that is starving.

4. MEASURES FOR REDUCING PRIVATE CAR USE

4.1. Classification

Several measures may conceivably reduce the gravity of the adverse car-use-related effects in metropolitan areas. Some of them (e.g., increasing the capacity of road infrastructure and building clean cars) do not necessarily require a reduction in car use. However, according to a general assessment of the current state, the sole short-term solution in urban areas is to reduce car use and to change use with regard to time and space, particularly on major commuter arteries during peak hours and in downtown districts.

TABLE I
Measures for Reducing Car Use

Prohibition or physical restriction of car traffic
Fuel rationing
Taxation of cars and fuel
Road or congestion pricing
Parking control
Decreasing speed limits
Not constructing new road infrastructure
Teleworking
Land use planning
Bus and high-occupancy vehicle lanes
Park and ride schemes
Improved transit
Improved infrastructure for walking and biking
Marketing of alternative travel modes
Public information campaigns
Social modeling

Widely proposed measures targeting reduced demand for private car use include increasing the cost of driving; prohibiting or physically restricting car traffic; improving alternative travel modes, such as transit, cycling, and walking; and changing the relative locations of homes, workplaces, shopping facilities, and recreational facilities so that driving distances are reduced. Table I lists several measures. These differ in effectiveness, cost, technical feasibility, and political feasibility.

Examples of implemented measures to manage travel demand are the prohibition of car traffic in downtown Cambridge, United Kingdom; road or congestion pricing in Singapore; and marketing of alternative travel modes in Perth, Australia. These three measures vary from more to less coercive—that is, in the degree to which adaptation is forced on car users.

4.2. Effects on Private Car Use

A continuum ranging from more to less coercive is obviously related to the effectiveness of measures for managing demand for car use. Implementing coercive measures is difficult, however. Despite awareness of the adverse effects of car use, car users in general are opposed to coercive measures since they impinge on their freedom to use the private car. Improving alternative travel modes is more popular. Also, coercive

measures are believed to be unfair since they affect everyone. At the same time, car users prefer more to less coercive measures since they minimize free riding and are believed to be more effective. At the political level, conflicts of interest make it difficult to reach consensus concerning the need to implement coercive measures. These difficulties are augmented by politicians' desire to win votes to remain in power if they believe (not always correctly) that their voters are not in favor of the measures.

The effectiveness of noncoercive measures depends on how able and willing car users are to reduce car use. Several moderating or mediating factors have been identified. Providing acceptable alternative travel modes is a basic requirement. However, car use has frequently developed into a habit so that information about alternatives is no longer sought. Important changes such as providing new alternatives may therefore go unnoticed. Information campaigns targeting the general public as well as individualized information about how to change travel have nevertheless had limited success.

Alternative travel modes are difficult if not impossible to make more attractive. Therefore, in addition to providing them, car use must be made less attractive. Introducing fees for car use is a preferred measure for doing this. In this way, some of the costs for improving alternative travel modes such as transit or constructing new infrastructure for walking and biking may be covered. However, substantial cost increases appear to be required to achieve sizeable changes in car use. Implementing new pricing measures or increasing existing ones are therefore neither accepted by the public nor politically feasible.

4.3. Consequences of Reducing Private Car Use

Private car use results from the interrelated choices made by car owners that are dependent on their physiological needs, obligations, or personal desires to participate in activities such as work, shopping, and leisure. Car-use reduction must therefore be viewed more broadly as an adaptation to changes in travel options that potentially have consequences for engagement in activities.

Research is beginning to disentangle the process of adaptation of car use. A well-founded hypothesis states that in this process households or individuals set car-use reduction goals and form plans to attain the goals. In selecting change options, they start with those that are psychologically less costly. Since less costly options are also likely to be less effective, the process entails a hierarchy of choices of options over time from less costly and effective to more costly and effective. A possible hierarchy is shown in Table II.

TABLE II
Adaptation to Reduced Car Use

Adaptation	*Psychological cost*
More efficient car use (trip chaining, car pooling, and choice of closer destinations)	Additional planning
Trip suppression	Activity suppression
Switching to alternative travel mode	Increased time pressure, inconveniences

If change is forced on car users, the costs of smaller changes may be less flexibility and, as a consequence, an increased need for planning. Additional changes may make necessary unwanted lifestyle changes that restrict the opportunity to perform some activities. Switching to alternative travel modes that are slower and less comfortable incurs the additional cost of increased time pressure.

4.4. Car-Use Reduction and Information Technology

Table III lists various devices that current information technology offers as components of intelligent transportation systems that are likely to decrease the psychological costs of adaptation to reduced car use. Information technology may make both substitution of travel feasible and car use more efficient. Examples of the former are flexible work arrangements, teleconferencing, and electronic commerce. Flexible working arrangements (e.g., telecommuting and a compressed workweek) lead to an overall reduction in car use due to reduced commuting. However, any savings are often offset by new, longer car trips for noncommuting purposes, many of which were previously linked to the commuting trips. Teleconferencing has the potential to decrease travel to the extent that the technology required is sufficiently developed. However, it is often the case that long trips, as opposed to short trips, are most likely to be substituted by teleconferencing. Electronic commerce also has the potential to suppress trips, primarily service and maintenance trips such as

TABLE III
Decision Support and Substitutes Provided by Information Technology

Adaptation	*Information technology application*
More efficient car use	Invehicle navigation systems
	Pretrip/enroute congestion or accident information
	Car pooling or car sharing information
	Information about road use charges
Suppressing trips	Teleworking
	Teleconferences
	Electronic commerce
Switching mode	Online public transport time tables and routing
	Information about road use charges
	Park-and-ride information
	Information about traffic regulations

for paying bills and other nongrocery shopping (e.g., book and music purchases), although inroads have been made with respect to grocery shopping.

Information technology making car use more efficient includes pretrip or en route information about travel time, shortest routes, congestion and accidents, and available parking spaces, as well as in-vehicle navigation systems. Information technology thus supports choices of departure times, routes, and destinations. Drivers are remarkably sensitive to such information, even when its role is not directly and explicitly relevant to route choice. Information technology may also make information better and more easily accessible with regard to alternative travel modes such as transit, thus possibly influencing travel mode choice. Furthermore, more efficient car use can be facilitated through the increased use of car-sharing schemes, whose efficiency and attractiveness (e.g., online bookings and payments) can be much enhanced with information technology.

Other information technology applications include electronically charging drivers when road pricing schemes are implemented. An essential aim of such applications is to provide information about the expected costs, particularly if the costs vary with zone, time of day, and congestion. Road pricing is of course not likely to affect car use unless such information is accessed, understood, and acted on. If these conditions are fulfilled, however, information about road pricing allows drivers to avoid congestion. Furthermore, providing cost information may facilitate not only the implementation of car-use reduction goals but also the setting of such goals.

The effectiveness of prohibiting car traffic from entering protected downtown areas may increase if coupled with easily accessible pretrip information about alternative travel modes and destinations (e.g., parking information or information about park-and-ride options). Such coupling may reduce the likelihood of car users driving as near as possible to the prohibited area instead of switching to transit or walking at an earlier point.

See Also the Following Articles

Aviation ■ Driving Safety ■ Transportation Systems, Overview

Further Reading

Ahrentze, T., & Timmermans, H. J. P. (2000). *ALBATROSS: A learning based transportation oriented simulation system*. Eindhoven, The Netherlands: European Institute of Retailing and Services Studies.

Axhausen, K., & Gärling, T. (1992). Activity-based approaches to travel analysis: Conceptual frameworks, models, and research problems. *Transport Reviews*, *12*, 323–341.

Bovy, P. H. L., & Stern, E. (1990). *Route choice: Wayfinding in transport networks*. Dordrecht, The Netherlands: Kluwer.

Button, K., & Hensher, D. (Eds.). (2001–2003). *Handbooks in transport*. Amsterdam: Pergamon/Elsevier.

Gärling, T., Eek, D., Loukopoulos, P., Fujii, S., Johansson-Stenman, O., Kitamura, R., Pendyala, R., & Vilhelmson, B. (2002). A conceptual analysis of the impact of travel demand management on private car use. *Transport Policy*, *9*, 59–70.

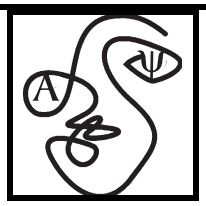
Gärling, T., Gärling, A., & Loukopoulos, P. (2002). Forecasting psychological consequences of car-use reduction: A challenge to an environmental psychology of transportation. *Applied Psychology: An International Review*, *51*, 90–106.

Golob, T. F. (2001). Travelbehaviour.com: Activity approaches to modeling the effects of information technology on personal travel behaviour. In D. Hensher (Ed.), *Travel behavior research: The leading edge* (pp. 145–183). Amsterdam: Pergamon.

Hanson, S. (Ed.). (1995). *The geography of urban transportation*. Amsterdam: Elsevier.

Novaco, R. (2001). Psychology of transportation. In N. J. Smelser, & P. B. Baltes (Eds.), *International encyclopedia of the social and behavioral sciences*. Amsterdam: Elsevier.

Schade, J., & Schlag, B. (Eds.). (2003). *Acceptability of transport pricing strategies*. Oxford, UK: Elsevier.



Trust

Ana Cristina Costa
Delft University of Technology, Delft, The Netherlands

1. Introduction
2. Defining and Measuring Trust
3. Developing Trust
4. The Benefits of Trust
5. Challenges and Dilemmas of Trust: Implications for Practice
6. Emerging Issues in Trust Research
7. Summary
Further Reading

GLOSSARY

calculus-based trust Trust based on calculation most likely to be found at the start of new relationships or partnerships when the parties do not have any prior connections or knowledge about each other.

identification-based trust Trust that arises from identification and empathy with the other party's desires and intentions and that results from sharing a common identity and similar values.

interpersonal trust Trust that is personalized and concerns trust between people that is characteristic of primary and small group relationships.

knowledge-based trust Trust that derives from knowledge built from prior interactions and/or collaborations allowing partners to have enough information to predict each other's likely behavior.

swift trust Trust that develops when people interact more in a role-based manner than in a person-based manner.

system or institutional trust Trust that is abstract and refers to trust in the functioning of organizations, institutions, political systems, and societies as a whole.

Trust is central to human life and is considered to be essential for stable relationships, fundamental for maintaining cooperation, vital to any exchange, and necessary for even the most routine of everyday interactions. In organizations, the importance of trust has been recognized at both the interpersonal and institutional levels. Two types of trust can be distinguished: interpersonal trust, which refers to trust between people, and system or institutional trust, which refers to trust in the functioning of organizational, institutional, and social systems.

1. INTRODUCTION

In organizational science, applied psychology, and related fields, trust has become a major focus of theory and research during recent years. Although scholars have long been interested in the study of trust, changes in the functioning of organizations and in scholarly thinking have led to a revival of interest in this topic.

Contributing to the rise of trust on the research agenda is the growing evidence of the wide and varied benefits of trust for individuals, teams, and organizations. Numerous studies have demonstrated how increases in trust result, either directly or indirectly, in more positive workplace behaviors and attitudes, better team processes, and superior levels of performance. Considerable efforts have also been expended to apply emerging trust theory to a variety of important organizational problems, with some being the

©2004 Elsevier Inc.
All rights reserved.

result of the increase of distrust in institutions' policies and management. Although trust might not be the ultimate solution for all problems, as organizations have moved toward flatter and more team-based forms of organized activity, interpersonal dynamics—and trust in particular—have become critical elements in achieving effective collaboration. Perhaps more now than ever before, organizations must invest in conditions that facilitate trust among members to survive and remain effective.

Trust becomes a vital concept when there are significant risks involved in trusting (i.e., vulnerability) and when there is objective uncertainty about future consequences of trusting. In organizations, uncertainty and vulnerability arise for different reasons. Consequently, trust has been studied with regard to different respects and has been approached through different perspectives. This article reviews the progress regarding the conceptualization and measurement of trust. It discusses how trust emerges and develops through different forms, and reflects on important benefits of trust at different levels of the organization. In addition, it discusses some of the most important challenges for research and practice in this area.

2. DEFINING AND MEASURING TRUST

Because trust is so central to human relations, many definitions have been put forward from a variety of perspectives, but to date there is no agreed-on general definition. In general, trust has been conceptualized as either a one-dimensional or a multidimensional phenomenon.

2.1. One-Dimensional Conceptualizations

One-dimensional conceptualizations include definitions of trust as a psychological state or as a behavioral choice. The psychological state approach defines trust in terms of interrelated cognitive processes and orientations toward beliefs or positive expectations in relation to others. Some definitions apply to contexts where these others are identifiable, for example, Rousseau and colleagues' definition of trust: "a psychological state comprising the intention to accept vulnerability based upon positive expectations of the intentions or behavior of another." In other definitions, trust is presented as a more general attitude or expectancy about other people or about the general social system. Common to these definitions is the reference to states of "vulnerability," "confidence," and "positive expectations."

Trust as a choice behavior can be seen as the willingness to take risks by acting on the basis of words, actions, or decisions of others. Some definitions portray trust as a more or less rational decision, motivated primarily by perception of risks, concerning either the probability of successful cooperation or the possibility of transaction cost reduction. More recent definitions suggest the need to conceptualize trust not only as a calculative orientation toward risk but also as a social orientation toward other people and societies as a whole. As recent research has demonstrated, one not only "thinks" but also "feels" trust. Throughout research various behaviors have appeared to be indicative of trust, including cooperative behaviors such as open communication, acceptance of influence, forbearance from opportunism, and lack of monitoring.

Despite the differences, trust as a psychological state and trust as a choice behavior are in principle compatible approaches. In most definitions, trust is related to individual attributions about other people's intentions and motives underlying their behavior. These attributions have cognitive and emotional grounds, and they influence and are influenced by individuals' general beliefs and expectations about the treatment they will receive from others. Finally, these attributions are closely linked to the decision to engage in behaviors of trust when interacting with others.

2.2. Multidimensional Conceptualizations

In discussing how expectations underlying trust affect subsequent behavior, several scholars have alluded to the fact that an adequate analysis of trust begins by recognizing its multifaceted character. Lewis and Weigert suggest that trust is a highly complex phenomenon containing distinct cognitive, emotional, and behavioral dimensions. People trust on the basis of emotions and cognitive processes that discriminate others as being trustworthy, untrustworthy, or unknown. The behavior reflects the significance of these earlier dimensions and enables individuals to act on their judgments. Mayer and colleagues argue that the assessment of one's trustworthiness influences and is influenced by a general propensity to trust others that develops from past experiences and

helps to shape the general beliefs about the treatment people expect to receive from others. Several authors have described trust as a function of different interrelated components, including individual propensities, interpersonal assessments of trustworthiness, behaviors of cooperation, and lack of monitoring. Gambetta explicitly connected positive expectations with cooperation: "When we say we trust someone or that someone is trustworthy, we implicitly mean that the probability that he will perform an action that is beneficial or at least not detrimental to us is high enough for us to consider engaging in some form of cooperation with him."

Besides defining trust in different ways, the approaches discussed here suggest that trust may be composed of different elements, each of a different nature. Individual predispositions, characteristics, and intentions of the trustee(s) and situational conditions such as the degree of risk can determine both the level and the potential form that trust takes. The relative importance of these factors is determined by the type and course of the relationships in where trust occurs.

2.3. Trust Measures

The conceptual diversity regarding trust is also reflected in the instruments developed to measure this concept in different contexts and levels of analysis. An overview of the most relevant measures of trust for organizational settings is given in Table I. These trust measures focus on different components and are based on different conceptualizations. For example, instruments of trust that assess another party's trustworthiness are often related to definitions of trust as a psychological state (e.g., Conditions of Trust Inventory [CTI]), whereas measures that emphasize behaviors are associated with behavioral choice definitions of trust. Multidimensional definitions of trust, on the other hand, include perceptions of both trustworthiness and trust behaviors (e.g., Organizational Trust Inventory [OTI], Team Trust Inventory [TTI]).

Multidimensional measures have the advantage of assessing simultaneously different facets of trust in a particular relationship context. However, one-dimensional measures can be useful for explaining differences between individuals or groups with respect to one of the trust facets. Instruments that measure the propensity to trust in general (e.g., Interpersonal Trust Scale [ITS], Revised Philosophies of Human Nature [RPHN]) can be used to explain differences in trust behavior between individuals or groups in the same situation. Some individuals tend to give most people the benefit of the doubt by trusting loved ones and strangers alike until experience shows that this trust is not warranted. Others expect only the worst of everyone around them. Although propensity to trust itself does not determine whether or not a person will trust in a specific situation, as situations become increasingly unfamiliar, its influence increases. In modern organizations, for instance, the growing need for cooperation between and within boundaries brings together people who do not necessarily know each other or have a previous history of working together. In such situations, the general willingness to trust others can strongly determine initial trust decisions.

Instruments that focus on trustworthiness and/or trust behaviors are situation specific and are particularly relevant in contexts of interpersonal and group relationships. Expectations about an individual's trustworthiness have been consistently found to be a strong determinant of trust in contexts of specific others. In general, the assessment of trustworthiness is based on three primary criteria—benevolence, competence, and integrity—the relative importance of which can differ from situation to situation. In some situations, competence can be more important. Other situations can demand political sensitiveness, making an individual's integrity more important.

Measuring trust behaviors in a particular context can be useful for learning about an individual's motives and intentions and for being able to make inferences of trustworthiness. Trust behaviors reflect the willingness to be vulnerable to others whose actions an individual does not control and can also differ across contexts. For instance, in strategic alliances, trust behaviors may include financial and formal investment, whereas within teams, cooperation and lack of monitoring would be more reflective of trust.

The measures described here assess different facets of trust, and most have limited applicability to particular contexts and levels of analysis. The choice of a particular measure over another implies some thought about the approach to trust chosen, the level of analysis in which trust is measured, and the purpose for which the instrument is used. In exploring differences between trusters, propensity measures can be more relevant. Differences between trustees can be better explained through measures that focus on trustworthiness. Behavioral measures of trust reflect the trust given in a particular setting. The whole process of trust in particular contexts can be better explored through multidimensional measures.

TABLE I
Overview of Trust Instruments Used in Organizational Contexts

Instrument/Author(s)	*Components of trust measured*	*Factors identified*	*Applicability/Level of analysis*
Interpersonal Trust Scale (Rotter, 1967)	General willingness to trust others	Trust in peers Trust in institutions	Assesses credibility of social agents and societal optimism Suitable to measure individual's trust in generalized others
Revised Philosophies of Human Nature (Wrightsman, 1964)	Propensity to trust others	Trust Cynicism	Measures expectations about the way in which people generally behave Indicates the general ability to trust others
Mayer & Davis (1999)	Propensity to trust Trustworthiness	Propensity Ability, benevolence, and integrity	Applicable for studying trust in organizations or top management
Cook & Wall (1986)	Trustworthiness	Faith in intentions and confidence in actions of peers and management	Applicable to study trust in peers and trust in management
Conditions of Trust Inventory (Butler, 1991)	Trustworthiness	Integrity, honesty, fairness, competence, consistency, loyalty, discreetness, openness, receptivity, availability, and fulfillment	Assesses trustworthiness of specific peers Is applicable to dyads in either horizontal or vertical work relationships
McAllister (1995)	Trustworthiness	Cognition based Affect based	Refers to managers' trust in peers Applies to dyads
Smith & Barclay (1997)	Trustworthiness Trust behaviors	Character, role competence, judgment, motives and intentions, relationship investment, acceptance of influence, communication, openness, and forbearance from opportunism	Distinguishes between perceived trustworthiness and trust behaviors in specific peers Applied in strategic alliances contexts
Currall & Judge (1995)	Trust behaviors	Communication openness, informal accord, task coordination, and surveillance	Applied to boundary role–persons dyads
Organizational Trust Inventory (Cummings & Bromiley, 1996)	Perceived trustworthiness Behavior intentions	Keeping commitments Honest in negotiations Does not take advantage	Multidimensional measure of trust Applicable to dyads or business units
Team Trust Inventory (Costa, 2003)	Propensity to trust Perceived trustworthiness Trust behaviors	Propensity to trust Perceived trustworthiness Cooperative behaviors Lack of monitoring	Multidimensional measure of trust for team contexts

3. DEVELOPING TRUST

It is a matter of common understanding that trust is not static. Trust develops over time and goes through various phases such as building, declining, and renewing. People start a new relationship with a certain level of trust, either higher or lower, depending on their own individual dispositions, past experiences, familiarity with or knowledge of the other party, contextual contingencies, and perceived incentives for cooperation. Three bases for developing trust in work relationships have been identified: calculus, knowledge, and identification.

3.1. Calculus-Based Trust

Calculus-based trust is likely to be found in relationships that are new and are formed between partners or team members who do not have any prior social connections. The basis for trust is the calculation of the benefits of being trusting and trustworthy against the threat that if trust is violated, one's reputation may be damaged by the other person's network of friends or colleagues. In other words, calculation underpins the belief that the actions of another party will be beneficial and reliable rather than the opposite. Trust in this sense consists of ensuring consistency of behavior, that is, that people will do what they say they will do. Making agreements, establishing the "rules of the game," and monitoring compliance can help to build this mode of trust. If others act predictably by complying with the agreements, they will build a reputation of trustworthiness on which trust can develop.

3.2. Knowledge-Based Trust

Knowledge-based trust is based on prior interactions and cooperation. It occurs when one has enough information about others to understand them and to be able to predict their likely behavior. Knowledge-based trust relies on information rather than deterrence. Trust develops largely as a function of the parties having a history of interaction that allows them to develop a generalized expectancy that each other's behavior is predictable and trustworthy. Here, trust develops through information about preferences, wishes, and behaviors of the other party that develop over time as a consequence of the parties having a history of interaction. Partners or team members are likely to develop common ways of thinking through their sharing of experiences and information. Two key processes are responsible for achieving this stage: regular communication and interaction.

3.3. Identification-Based Trust

Identification-based trust is based on identification with the other party's desires and intentions. This form of trust arises between people who share a common identity, meaning that they hold similar values, including a shared concept of moral obligation. Trust develops through empathy because the parties effectively understand and appreciate the each other's wants. This mutual understanding is developed to the point where each party can effectively act for the other. Identification-based trust develops as both parties know and predict each other's needs, choices, and preferences and also share some of those as their own. Increased identification enables people to empathize strongly with the other party and incorporate parts of the other party into their own identity as collective identity.

These three basic conditions promote the emergence of trust and influence the individual expectations about the parties' trustworthiness and their willingness to engage in trusting behaviors and, consequently, their initial level of trust. If trust must be built from scratch, making agreements and monitoring compliance of team members can help to build trust. In mature relationships, trust may so solid, and the risks involved may be so small, that monitoring is not needed to maintain optimal cooperation. Calculus, knowledge, and identification can also be seen as sequential linked stages of trust where achieving trust at one level enables the development of trust at the next level. Although not all relationships develop fully and so might not develop past the first or second stage, effective cooperation is unlikely to take place or persist if relational trust does not develop between individuals who interact intensively over a period of time. Therefore, the development from calculus-based trust to knowledge-based trust is crucial.

4. THE BENEFITS OF TRUST

As organizations have come to rely less on structures and formal arrangements and more on collaboration and cooperation inside and outside the firms, trust is given more emphasis as one of the fundamental motors of these processes. If trust is absent, no one will risk

moving first and everyone will sacrifice the gains of collaboration and cooperation. Although laws, rules, and contracts are still necessary conditions for the stability and prosperity of organizations, for these to prevail, they must be based on reciprocity, moral obligation, and trust. The benefits of trust have been discussed in relation to reduction of transaction costs, collaboration, and cooperation within and between organizations as well as to the effects on performance and effectiveness.

4.1. Transaction Costs Reduction

Currently, competitive success has become increasingly dependent on the reduction of transaction costs as the requirements for quality have escalated internationally and markets have become more uncertain. Traditional transaction costs theory has neglected trust in its assumption that the risk of opportunism is high in the "governance" of relations. However, as many have recognized, trust is an element of every transaction that can be accounted for either by previous experience or by lack of contrary evidence. Moreover, the presence of trust in transactions is likely to generate more trust at other levels because transactions are embedded in professional and social networks, diminishing the hazard of opportunism. This does not necessarily mean that trust is a by-product of all transactions. The possibility of opportunism does exist. But, in the long run, opportunism can be very costly because it will increase the amount of control costs and most certainly will inhibit future transactions.

Reducing transaction costs through trust can be achieved through reputation. Having a trustworthy reputation can be of great "economic value" because it plays an important role in determining the willingness to enter into a business exchange with a given actor. This good repute will further lead to positive expectations in the future, enhance the level of trust, and promote the actor's willingness to cooperate. It has been argued that trust, besides reducing transaction costs and monitoring performance costs, also eliminates the need to install control systems that are designed to obtain short-term financial results. Nevertheless, trust should not be seen as a replacement for either market or hierarchical forms of transactions. These governance mechanisms are necessary for the establishment of communities, which can be seen as an important starter for trust based on shared ethical norms and values underlying the communities.

4.2. Collaboration and Cooperation within and between Organizations

It is commonly assumed that some level of trust must exist so that cooperation between partners can be achieved freely. Yet, trust as a precondition to cooperation can be subjected to different demands of intensity. Requirements for trust in organizations are dependent on the mechanisms that govern the cooperative decisions and the social arrangements in which those decisions are made.

Within interorganizational forms of collaboration, trust is important to the extent that it facilitates information exchange and reciprocity between partners, leading to organizational learning. When interfirm collaborations are forged from common memberships to a professional community, existing ties, or local community partnerships, trust seems to develop more on interpersonal bases and community networks. When alliances are forged mainly from mutual dependencies and/or calculation of resource needs, trust tends to develop from formal bases that can be more costly and time consuming. One way to facilitate collaboration between organizations is to make trust a part of the organizations' routines and practices so that collaborations between firms can continue successfully. This means that the major source of trust should be institutional. Key individuals or groups (i.e., boundary spanners) do play an important role in interfirm forms of collaboration. However, problems of turnover and the possibility of communication breakdown on the part of these individuals make trust at this level a very fragile form of governance.

Also within organizations, the importance of trust is recognized both at the institutional and interpersonal levels. Trust based on institutional arrangements, such as laws, rules, and professional practices that support the organization as a whole, create a common ground for understanding actions and enhance patterns of behaviors that are extendable to all organizational members. These create a general climate in which trust is produced and generalized to other levels. Although trust is important to the functioning of all forms of the organization, alternative forms have clear trust requirements and managerial philosophies have clear implicit levels of trust. Failures to meet these requirements bring various consequences. For instance, in functional organizational forms, insufficiency of trust reduces efficiency; in divisional forms, it reduces effectiveness and increases costs; in matrix forms, it causes the form to fail; and in networks, it

causes organizations to fail. The tendency toward flatter and more team-based forms of organized activity shows that the importance of trust in organizations has been augmented significantly. In the network form, it has become one of the requirements for organizations' survival.

New policies emphasizing interpersonal and intergroup dynamics in the workplace have accentuated the importance of trust at an interpersonal level. Interpersonal trust is a product of rational decisions and emotional bonds and can be based on different mechanisms, depending on the degree of knowledge or familiarity among the people involved. In situations where individuals have accumulated meaningful knowledge and have established some kind of bond with one another, interpersonal trust tends to be more based on the attributions that individuals make about the other person's character and the motives and intentions underlying these actions. In situations where individuals have little information about one another or have not yet established any kind of bond with one another, trust may initially develop on the basis of individual dispositions, situational constraints, and/or institutional arrangements.

4.3. Effects on Performance and Effectiveness

Apart from the general assumption that trust is an important lubricant of the social system and a facilitator of coordinated action among individuals, several important benefits have been associated with trust. Trust has been associated with positive work attitudes and behaviors such as open communication and information exchange, acceptance of influence, citizenship behavior, commitment, satisfaction, and reduction of conflict. Trust has also been associated with the ability to enhance collaboration and mutual learning between individuals by leading to cooperation and higher individual and team performance.

Some authors have labeled these effects as spontaneous sociability. When operationalized in behavioral terms, spontaneous sociability refers to numberless forms of cooperative, altruistic, and extra role behavior in which members engage that enhance collective well-being and further attainment of collective goals. It should be noted that there is also evidence showing that trust alone is not always enough and that spontaneous sociability also depends on the individual's perception regarding the efficacy of his or her own actions. In addition, implications of trust for performance effectiveness should be interpreted carefully. First, trust cannot be seen as one of the main indicators of performance given that tasks require specific abilities and knowledge to be performed adequately. In situations where individuals or teams do not possess adequate skills and knowledge to accomplish their tasks successfully, trust probably will not improve performance. Second, performance has been found to be dependent on numerous determinants, meaning that trust is just one of these indicators. In certain conditions, however, trust may play a more a moderated role by facilitating communication and openness, which can lead to the exchange of important knowledge or generate critical discussions that may be beneficial for the end product. The importance of such effects is again dependent on the trust requirements associated with the functioning of teams and organizations.

5. CHALLENGES AND DILEMMAS OF TRUST: IMPLICATIONS FOR PRACTICE

Desirable development and maintenance of trust, however, is neither easy nor ensured. Newspaper headlines during recent years have headed stories about the collapse or fraud of several major corporations. Although reorganization processes within companies are essential for survival in the current volatile environment, they often have bee accompanied by a considerable lack of trust between organizational members. Ironically, the increased need for trust in modern organizations has also made the role of checks and controls even more important. This is because violations of trust are more likely to occur when vulnerability increases, because excesses of trust can lead to miscalculations of risks or illusions of greatness. In either case, such effects can lead to drastic consequences for organizations. To deal with some of the challenges and dilemmas of trust, it is important to recognize two important issues regarding trust in organizational settings: the fragility of trust and the limitations of trust.

5.1. Fragility of Trust

It is often suggested that trust is typically created rather slowly, but it can be destroyed in an instant by a single mishap or mistake. One of the reasons for this is that trust is a phenomenon that feeds on itself. When a relationship or partnership is growing, trust builds on

evidence of trustworthiness and on a track record of being trustworthy. If trust-building actions are taken, the overall level of trust grows until it begins to even out during the maintenance stage. The building process is often slow because people tend to be reticent about trusting, particularly with regard to those who are not known. Once trust has been built, the demands for evidence of trustworthiness diminish, and this can lead to a false sense of security. In some cases, parties may actively discount information that implies untrustworthiness, at least for a while. However, if solid evidence of untrustworthiness emerges, trust is destroyed quickly and the collapse is dramatic.

Both psychological and structural processes can explain the asymmetry between creating trust and destroying trust. Because people tend to pay more attention to negative events (trust destroying) than to positive events (trust building), negative events carry greater weight in people's trust judgments. This is partly because negative events are more visible and frequently take the form of specific, well-defined incidents such as lies and discoveries of errors, whereas positive events, although sometimes visible, are more often fuzzy or indistinct. The accumulation of relevant experience with low trust can frame the interpretation of events to reinforce this belief.

Organizational structures and management policies also contribute to the fragility of trust in organizations. By becoming less bureaucratic and more focused on collaborative processes, organizations have increased the level of uncertainty and vulnerability between their members. Trust based only on formal mechanisms is insufficient for organizations to function effectively. In modern organizations, both institutional trust and interpersonal trust constitute an essential feature of effective functioning. Successful companies or work relationships develop, treasure, preserve, and nurture trust at both levels. They recognize that trust helps during both good and bad times, and they recognize that if trust is lost it might never be recovered.

5.2. Limitations of Trust

Trust is pervasive and indispensable to the functioning of organizations, but when it is excessive or when its limitations are not recognized, trust might not be prudent and can have adverse impacts. Because trust goes beyond the available evidence, there is always an amount of unknown risk that results from the lack of control or of complete knowledge about future outcomes and about the actions of the other party.

Excesses of trust can be harmful when they entail a large risk of economic damage and threaten the survival of organizations. From interpersonal to organizational interaction, excesses of trust can have adverse impacts on the functioning of firms. For instance, a culture of excessive trust and benevolence may become oppressive when it results in the expression of criticism being taboo and when it blocks the direct voicing of complaints, thereby preventing fast and direct solutions between partners in conflicts. It can yield "groupthink," which reflects excessive cohesiveness and a shared illusion of invulnerability. This can lead to an accumulation of unsolved conflicts, biasing perceptions of evidence of incompetence or malevolence.

Excesses of trust can occur for different reasons. One is naivety, ignorance, or cognitive immaturity that makes one party unaware of risks involved in the trust situation. Here, the experience of broken trust will teach awareness. Another reason is the feeling of omnipotence, that is, an overestimation of one's power to control an untrustworthy partner with the feeling that one cannot be hurt by damage imposed by the partner. Another reason is impulsiveness deriving from carelessness or from putting a large emphasis on the current benefits relative to later adverse effects. This can be related to greed.

5.3. Implications for Practice

Recognition of the importance of trust for the functioning of organizations has grown considerably. Although the success of some organizations and work relationships can be more dependent on trust than others, in general, those organizations and work relationships that are successful find ways in which to build and maintain trust. Some are built on stronger foundations than are others, but the management of trust is always critical.

Leaders and managers play a central role in managing the overall level of trust as well as more specific levels of trust across organizations. Because leaders and managers are responsible for combining strategy, structure, overall operating logic, resource allocation, and governance of organizations, the levels of trust displayed by these mechanisms might well be reciprocated at other levels of organizations. When managers are perceived as competent, fair, and open regarding the sharing of information, the overall level of trust in organizations is likely to improve. Leaders determine, or at least influence, the degree of interdependence between organizational units or individuals, and this in turn determines the level of trust reliance between parties that is necessary to achieve

the work goals. Although trust does not necessarily leads to cooperation, interdependence across units, teams, or individuals creates conditions for information exchange and proximity, and these may allow room for acceptance and trust development. Managers can also try to create a climate that supports trustworthy behaviors between members and, at the same time, provides adequate guidance so that trust does not become dangerous.

The increasing emphasis on short-term and highly interdependent multiproject environments brings together people who are not necessarily part of the structure, culture, and norms of the same organization. Also in such contexts, developing trust between individuals or partners becomes essential. Lacking a common ground, a certain degree of familiarity, and prior experience in working together can influence the basis on which trust might develop. Recent research has described the so-called "swift trust" as a new form of trust that develops quickly, mostly in contexts where there is no previous history of interaction between partners, there is no prospect of long-term interaction, and people interact more in a role-based manner than in a person-based manner. The concept of swift trust suggests that members initially import trust rather than develop trust. Because there is insufficient time for these expectations to be built from scratch, they tend to be imported from other settings and imposed quickly in categorical forms. The categories invoked to speed up these perceptions can reflect roles, industry recipes, cultural cues, and occupational and identity-based stereotypes. However, this form of trust might not be enough to function effectively. For trust to develop further, parties should act trustworthy by complying with the agreements and expectations. Yet the short life span of some of the current work relationships also increases the opportunity to take advantage. Again, the involvement of supervisors and managers can be crucial in maintaining trust at an optimal level that benefits all parties involved.

Trust is a significant decision-making process under conditions of vulnerability and uncertainty. Managing trust reflects awareness of the mental processes behind the decision to trust and awareness of the risks involved in such a decision. Risk assessment involves considerations about another party's trustworthiness and about the situational factors that weigh the likelihood of positive and negative long-term effects of the trust. Because the information about another party's competence, motives, and intentions is often incomplete, people recall past experiences as their best guides. In more unfamiliar contexts, these considerations can be more dependent on individual propensities, reputation, or external information by inquiring about another party's trustworthiness. Assessing the risk before trusting is crucial to this process because when trust is not fulfilled, the trusting party suffers an unpleasant consequence that is greater than the gain that the trusting party would have received if trust had been fulfilled. Moreover, if trust is broken, it might be destroyed completely or might require a long time to be rebuilt to its former state. Therefore, trust must be preserved and nurtured by information that reinforces perceptions of trustworthiness and trustworthy behavior.

6. EMERGING ISSUES IN TRUST RESEARCH

Despite the considerable development in theory building regarding trust in organizations, issues concerning the definition and its measurement at various levels of analysis remain the subject of continuous discussion. The growing globalization of business has increased the emphasis on collaboration and cooperation across organizations and even countries, raising the issue of trust across cultural borders. How trust develops between representatives from various organizations and/or countries, what basis of trust will be sustained in such contexts, and what factors (both personal and organizational) will determine the level of trust must be explored more comprehensibly. Similarly, the increasing competition of markets has led many competing firms to form alliances to survive. Whether trust in such contexts can develop from a calculative basis into more mature stages to achieve effective cooperation and whether partners will trust one another with their business secrets to promote learning are questions that need to be researched in more detail. In addition, the rapid advances in information and communication technology have created virtual work environments. Understanding the role of trust in such environments, how trust might develop without face-to-face interaction, and what factors are considered to assess another party's trustworthiness has become extremely important.

7. SUMMARY

Trust is a central to human life as well as organizational life and is one of the pillars on which organizations function. Despite the numerous benefits of trust and the current high value placed on it, trust also has

limitations. This article has focused on the definition of trust, on issues concerning the development and maintenance of trust, on the benefits of trust in organizational settings, and on challenges and dilemmas associated with trust. Emerging issues in trust research were noted, including the development of trust in a diverse workforce and across organizational borders, in competitive environments, and in virtual environments.

See Also the Following Articles

Cooperation at Work ■ Leadership and Culture

Further Reading

Costa, A. C., Roe, R. A., & Taillieu, T. (2001). Trust within teams: The relation with performance effectiveness. *European Journal of Work and Organisational Psychology*, *10*, 225–244.

Dirks, K. T., & Ferrin, D. L. (2001). The role of trust in organizational settings. *Organization Science*, *12*, 450–467.

Gambetta, D. (Ed.). (1988). *Trust: Making and breaking cooperative relations*. Oxford, UK: Basil Blackwell.

Kramer, R. M. (1999). Trust and distrust in organizations: Emerging perspectives, enduring questions. *Annual Review of Psychology*, *50*, 569–598.

Kramer, R. M., & Tyler, T. R. (1996). *Trust in organizations: Frontiers of theory and research*. Thousand Oaks, CA: Sage.

Lewis, J. D., & Weigert, A. (1985). Trust as a social reality. *Social Forces*, *63*, 967–985.

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *Academy of Management Review*, *20*, 709–734.

McAllister, D. J. (1995). Affect- and cognition-based trust as foundations for interpersonal cooperation in organizations. *Academy of Management Journal*, *38*, 24–59.

Nooteboom, B., & Six, F. (Eds.). (2003). *The trust process in organizations*. Cheltenham, UK: Edward Elgar.

Rousseau, M. T., Stikin, S. B., Burt, S. B., & Carmerer, C. (1998). Not so different after all: Across-discipline view of trust. *Academy of Management Review*, *23*, 393–404.

Williamson, O. E. (1975). *Market and hierarchies*. New York: Free Press.

Urban Environments and Human Behavior

Gabriel Moser
Université René Descartes, Paris, France

1. Introduction
2. The Urban Environment as a Stressful Milieu
3. The Urban Environment as a Place of Living
4. The Rationale of Intervention in Urban Environments
5. Public/Private Urban Environments: The Neighborhood
6. Urban Public Environments: The City
7. Toward Sustainability of Urban Environments
Further Reading

GLOSSARY

appropriation A particular affective relationship to an object, which then becomes part of oneself.
attachment An emotional bond, which people often develop to their life space.
people–environment congruity The relation to a living place that corresponds to the aspirations of the individual and allows him or her to satisfy his or her material and social needs.
place identity Either a person's expressed identification to place or an aspect of his or her identity comparable to social identity.
socioenvironmental engineering Intervening on the physical as well as the social aspects of an environment involving all the concerned actors.
sustainability A development capable of satisfying the needs of the current generation without compromising the ability of future generations to satisfy their own needs.
urban behavior Any specific protective behavior adopted by city dwellers to cope with exposure to urban stressors.
urban stressors The physical (noise and pollution) or social (density) environmental conditions that require various forms of coping.

It is estimated that by the year 2015 approximately half of the world's population will be living in large cities. Cities are characterized by their cultural, religious, ethnic, economic, and social diversity and their stressful living conditions. Citizens are exposed daily to a variety of environmental stressors, such as noise, pollution, or density, which make them adopt specific adjustment behaviors characterized by unfriendliness and egoism. Environmental psychology is playing an increasing role in helping shape the urban environment in order to conserve or create well-being for citizens. Socioenvironmental engineering acts on both the physical and the social aspects of the environment. It is mainly focused on the neighborhood but often extends to the relationship between the individual and the urban structure as a whole. Socioenvironmental engineering is a way to meet people's requirements for well-being and satisfaction and to achieve urban sustainability.

1. INTRODUCTION

Since they concentrate employment opportunities, culture, entertainment, and recreation, cities have

©2004 Elsevier Inc.
All rights reserved.

traditionally been a pole of attraction. It is estimated that approximately half of the world's population lives in large cities. Cities have progressively become the "natural" habitat of most individuals.

Although for some inhabitants the urban environment provides an ideal level of stimulation, an optimal context for their activities, and numerous opportunities for interpersonal contacts, urban daily life conditions have become increasingly difficult. How do these reputedly bad environmental conditions affect urbanites' behavior? When examining human behavior in urban environments, the following aspects have to be taken into account:

1. The urban environment as a stressful milieu (i.e., a place of daily exposure to a variety of specific stressful environmental conditions)
2. The city as place of living (i.e., appropriated by the individual and to which he or she identifies)

Before analyzing how environmental psychology addresses life in cities, an overview of the previously mentioned aspects is provided with regard to urban behavior. How can environmental psychology help identify the processes that regulate and mediate the individual's relationship to his or her environment? The requirements of sustainability give a new impetus to environmental psychology by emphasizing individual needs and well-being. By dealing with the individual–environment relationship at different spatial levels (i.e., from dwelling to the neighborhood and the city), environmental psychology provides analyses and interventions aimed at improving urbanites' daily life at these different levels. The point of departure of this analysis is often the set of physical specifications of the environment (e.g., noise, pollution, and planning and layout of physical space) that act directly on the individual or are mediated by the social variables of the environment (e.g., crowding and population heterogeneity). However, the physical and social factors are inextricably linked with regard to their effects on individuals' perceptions and behavior.

2. THE URBAN ENVIRONMENT AS A STRESSFUL MILIEU

Exposure to the various stressors and particular life conditions of urban environments results in specific urban behaviors.

2.1. Urban Stressors

Urban environments display three specific aspects: ambient physical conditions (essentially noise and air pollution), social conditions (density), and overstimulation (exposure to a high number and variety of visual and other stimulations). Cities are more noisy, more dense, and more polluted.

Noise is the most common stress to which city dwellers are exposed on a daily basis. The levels of noise are associated with the size of the population. Noise is the most frequently mentioned form of stress and the one that leads to the most frequent complaints. One-fourth of urbanites are exposed to at least one loud noise at work, home, or during commuting on a daily basis.

Because of traffic, exposure to pollution is more frequent in large cities. Although the effects of air pollution on health (e.g., respiratory problems for children and elderly) are well documented, pollution has little direct effect on the behavior of urban residents.

Generally, people are more likely to perceive environmental problems when they can hear (noise), see (smoke), smell, or feel them. The relationship between exposure to air pollution and health is mediated by the perception of exposure. An important source of information is the media's interpretation of noise or pollution levels, which may influence public perceptions and attitudes.

Cities often have different forms of density due to large population concentrations. City dwellers may be exposed to social density outdoors (high number of individuals in one area) and to spatial density in their habitat (small amount of space per person) and, therefore, struggle for resources due to a low ratio of population to the amount of available resources (i.e., too many people and insufficient transport facilities, shops, and services). Furthermore, the large amount of visual and auditory stimuli to which city dwellers are exposed is a constant informational overload that increases fatigue.

2.2. Urban Living Conditions

Citizens not only are exposed to environmental stressors but also deal with poor living conditions. Mobility in cities is more constrained than that in small towns: Commuting daily from residence to workplace and dealing with frequent gridlock or crowded mass transit increases urbanites' stress. They have less spare time

because of longer commuting times, and there is less time available for social interactions.

Bureaucratic formalism is also more frequent in large cities than in small towns. Compared to inhabitants of small towns, urbanites experience more frequent difficulties accessing various services due to increased competition (subway, taxis, etc.).

The level of crime is higher in large cities. Therefore, urbanites are more likely to be exposed to violence and crime. The delinquency rate is higher in city centers than in suburbs, which is commonly explained by the numerous opportunities to commit crime in the city along with a lower probability of being caught. Fear of crime (which is not necessarily correlated with objective crime rates) restricts urbanites' behavior by making them feel vulnerable. However, despite being generally more exposed to environmental stressors, the inhabitants of large cities, with the exception of newcomers, do not cite these factors as annoying them more frequently than do the inhabitants of small towns or even people living in rural habitats.

Overall, city life seems to be constraining and demanding. All analyses of environmental overload, environmental stress, and behavioral constraints note the potentially negative effects of life in large cities compared to life in small towns. When exposed to such stressful situations, city dwellers engage in coping processes. Coping is an attempt to reestablish or gain control over a situation regarded as stressful. Therefore, in order to achieve congruity between themselves and the environment, urbanites adopt protective adaptation processes and perform behavioral and/or cognitive changes, which engender specific urban behavioral norms.

2.3. Specific Urban Behavior

At the beginning of the 20th century, Simmel noted the social withdrawal and egoistic behavior of urbanites, as well as their detachment and disinterest toward others. In 1970, Milgram outlined three urban adaptation behaviors: coping by choosing priorities, self-protection by erecting psychological barriers around oneself, and creating rules and institutions.

Urban behavioral particularities comprise the following: segmented and functional ways of interacting with one another, anonymity and lack of involvement, indifference toward deviant and bizarre behaviors, and restriction and selectivity of responses to other peoples' demands. Environmental constraints such as gridlock, overcrowding, and fear of crime enhance behaviors aimed at gaining control and freedom of action. The pace of life is more hectic. Certain activities are performed more quickly in large cities than elsewhere. People walk faster in large cities than in small towns. The fear of crime provokes behavioral restrictions, especially among elderly people. Withdrawal and decreasing responsiveness to the diverse demands of other people are typical reactions to stress. People filter inputs, focus attention on the most important demands, and neglect the peripheral stimuli, thus avoiding unwanted interactions. By concentrating on their own needs, urbanites not only deliberately avoid contact with others but also pay less attention to what is going on around them. The constant expression of such adjustment behavior in urban setting indicates that this is a normative behavior.

In terms of interpersonal relationships, individuals establish interaction priorities, excluding a certain number of contacts. In urban settings, interpersonal relationships are governed by rigorous rules, which enable individuals to preserve the minimum privacy needed in order to protect themselves from intrusion by others despite high-density situations. Gatherings with friends are more often planned in advance, and informal encounters are rare and limited to the home or workplace.

The characteristics of interpersonal exchanges in the urban environment not only tend to make social contacts more superficial and eliminate interactions that have no personal benefit to the individual but also lead to neglect expressions of politeness. Civility is a disinterested act. It involves a common code of conduct, which is indispensable for maintaining the social fabric, based on respect for the other. The individual holds the door open for others or says "hello" when entering a shop less often in large than in small cities.

In short, public spaces are locations for expressing community values, rules, and codes that are indispensable for individuals to maintain social distance and to protect their personal space. Social densities of large cities create both physical proximity and social distance. Thus, the regulation of social life in a densely populated society is based on adhering to routines that can be anticipated and predicted. Civil behaviors are indispensable for social interaction in the urban setting. On the other hand, incivilities may be regarded as the negation of the processes of the civilization of morals or as standards of behavior—that is, as breaches of social mores.

3. THE URBAN ENVIRONMENT AS A PLACE OF LIVING

People identify with the place where they live: "My home is my castle." Feeling at home is an important aspect of one's well-being and this is a universal human experience. Individuals create privileged relations to their habitats, which become their "homes," and their homes provide identity. This involves a way of building one's life within a geographically limited space. Several characteristics transform a habitat into one's "home": centrality, continuity, privacy, expression of self and identity, social relationships, atmosphere (warmth and pleasantness), and characteristics of the physical environment.

3.1. Place Identity and Attachment

There are two ways of relating one's place and one's identity. The first is place identification. This concept refers to a person's expressed identification with a place, which becomes part of his or her social identity. The second way of relating one's place and one's identity is through place identity, a specific aspect of the individual's identity.

An important mechanism that supports place identity is the attachment to a specific place. Individuals often develop an emotional bond to their life space, essentially their home and the neighborhood, but often also urban places and spaces on a larger scale. Such anchoring is an ongoing process, dependent on individual time perspectives: The duration of residence is essential for the individual's appropriation of his or her life space, which in turn is indispensable for well-being. Anchoring reflects the individual's motivations, social status, family situation, and projects for the future. Urban environmental appropriation revolves around forming social and interpersonal relationships. Individuals who make emotional investments in their neighborhood are more satisfied with their interpersonal relations in that neighborhood and develop a sense of well-being.

Appropriation may be regarded as a particular affective relation to an object that may then become part of the identity of the individual. Appropriation means having control over one's living space, and it is a prerequisite of feeling "safe" and "at home": It is essential to the construction of spatial identity. Appropriation is important for the individual to be able to organize and personalize his or her life space. This may be crucial not only in one's home but also in any other urban place in which one makes a temporal investment. Steady or transitionally occupied places produce place attachment and are often accompanied by ties to personal objects such as furniture, pictures and souvenirs, which indicate the appropriation of places.

Residential history determines how one appropriates home environments. The temporal horizon and needs that reflect one's position in one's life cycle affect one's ability to generate or failure to develop positive investments in the neighborhood. Finally, one can assume that individuals who appropriate their environment and feel at home where they live may also care more about the environment in general (i.e., more frequently exhibit ecologically beneficial behavior).

3.2. From Appropriation to Urban Identity

Appropriation operates not only at the level of one's place of dwelling but also extends to urban places, such as the street, the district, the town, or even the country, and it is accompanied by social networking (family, friends, neighbors, and communities). The urban identity is essentially acquired through various territorially bonded social networks. The feeling of being at home in one's neighborhood is linked to the frequency of encounters, the nature of local relationships, and the satisfaction that they provide. It involves social integration extended to local service providers such as physicians, shopkeepers, and others, and it constitutes the framework for the different individual networks (workplace, leisure, school, etc.). Furthermore, with regard to feelings of attachment, the social relations provided by a place may be more relevant than the place itself. Taking root corresponds to a desire for stability and permanency in one's way of relating to a certain place and one's involvement with the place in the long term. The sociospatial aspects of traditional urban structures include the residential environment—the district delimited by architectural, social, and administrative boundaries. The monofunctionalism of city structures has extended the way of relating to the city as a place of daily life beyond the traditional local district.

Increased residential mobility has resulted in a shift in place investment from one's housing to one's furniture and other "belongings" that contribute to the individual's identity. Moreover, settlement identity, which refers to individual preferences for certain

types of habitat, allows mobile individuals to conserve coherence and identity across various residences.

4. THE RATIONALE OF INTERVENTION IN URBAN ENVIRONMENTS

Environmental psychology analyzes individuals' and communities' perceptions, attitudes, and behaviors with regard to their explicit relationships with the physical and social contexts in which these individuals and communities live. Depending on the level of analyses, different aspects and problems are addressed.

4.1. Levels of Analyses from the Dwelling to the City

In dealing with the relationship between an individual and his or her life space, environmental psychology aims to understand and act in order to fulfill people's needs and aspirations. The notions of space and place are central. For each level of spatial reference, the discipline enables the investigation of people–environment interactions, whether involving an individual, group, or society (Table I).

The interactions and/or transactions between people and their urban environment at the different environmental levels may be understood by considering the individuals' perceptions, needs, opportunities, and means of control. In addition to appropriation and anchoring, control over one's environment plays a crucial role. Here, the scale of analysis allows one to account for a significant dimension in the individual–environment relation in terms of possibilities of control and mastery over the environment in question. Such potential control allows the individuals' aspirations to more or less master different aspects of their environment. These are important, individual, and direct with respect to the microenvironment. In one's habitat, appropriation is not shared with strangers, and the individual has full control of this place. In proximate and semipublic environments, the possibility of control and mastery is no longer individual but shared with a community. Participation at the local level is possible, and a sense of belonging may be created in certain circumstances, but control is necessarily mediated by other individuals. At the city level, control is beyond the range of the individual, and it may only be collective or societal. The possibility to control or master one's environment is an important requisite in anchoring and identity processes. How people relate to their place of living affects their well-being.

4.2. Socioenvironmental Engineering in Urban Environments

Environmental psychologists intervene at the interface of the physical and social aspects of the urban environment and engage in interdisciplinary problem solving in collaboration with architects and planners.

4.2.1. The Sociophysical Interface

Environmental psychologists intervene as socioenvironmental engineers, which means that they intervene as much on the social as on the physical aspect of the urban environment, examining the interrelation of these two aspects. They constantly bear in mind that an intervention on one of the aspects of the urban environment always has an impact on all the other aspects of the environment (i.e., changes in the physical aspect affect the social aspect and vice versa).

TABLE I
Social and Physical Aspects of Urban Environmental Extensions and Type of Control

	Physical aspects	*Social aspects*	*Type of control*
Level 1: Micro environments	*Private/semiprivate spaces:* Habitat, housing, building	Individuals/family	Extended control (unshared)
Level 2: Proximate environments	*Proximate public spaces:* Blocks of apartments, neighborhood, district	Neighborhood community	Shared control built on consensus
Level 3: Public urban environments	*Public spaces:* Town, city	Aggregates of individuals	Mediated control (participation possible at local level)

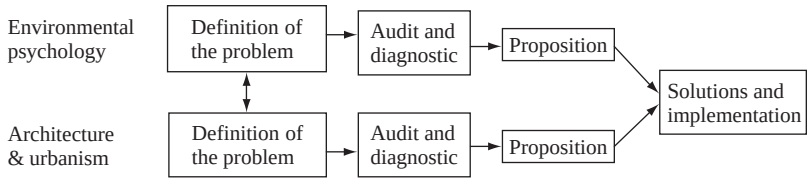


FIGURE 1 Rationale of interdisciplinary intervention.

Thus, environmental psychologists focus on the user of the environment as much as on the environment. Furthermore, each level of spatial reference has different implications that must be taken into account for any intervention.

Interventions in terms of socioenvironmental management concern public urban spaces and/or neighborhoods, and they consist of the diagnostic, audit, and intervention followed by a control after implementation of the intervention. By definition, they are implemented on existing structures. Socioenvironmental engineering also contributes to the understanding of the discrepancies between public environmental policies and individual aspirations, expectations, and beliefs. This may be important, for instance, with regard to noise or air pollution, but it also applies to the layout of the built urban environment. By identifying individuals' perceptions and needs, people–environment congruity can be enhanced.

4.2.2. Cooperation with Other Disciplines

The study of the urban environment requires multidisciplinarity, which is indispensable for environmental intervention. Responding to social demands often involves multiple approaches from different disciplines, each having its own characteristics and following its own disciplinary logic.

In urban matters, environmental psychologists often work in cooperation with members of other disciplines and professions, such as architects, engineers, landscape architects, and planners. Their respective analyses and interventions often have a different point of departure, focusing either on the built environment or on the people who live in it. Sometimes, a sequential gap may also appear: The architects, designers, and planners intervene at the conception and building process of the contexts in which people live, whereas social sciences typically focus on the existing structures.

Only knowledge integrated from the different disciplinary approaches may provide an adequate response to the challenges of living in cities. This integration has to follow a specific rationale in order to be efficient (Fig. 1).

The interdisciplinary approach is a multiple and parallel approach to the same problem, where each discipline provides a solution according to its own scientific logic that is then integrated in the global response. The cooperation between social sciences and design disciplines usually results in successful problem solving of urban issues because it allows all the different aspects of people's relations to the built environment to be addressed.

5. PUBLIC/PRIVATE URBAN ENVIRONMENTS: THE NEIGHBORHOOD

Interventions at the proximate level deal with the immediate environment of the individual's living space. This may be a row of houses or apartment building, the immediate neighborhood, or the leisure areas in the immediate surroundings of homes, such as parks and green areas. These areas are referred to as semipublic spaces, which means that control over them is shared within a community (inhabitants of a neighborhood, users, etc.).

5.1. People–Environment Congruity: Feeling at Home

Concepts such as "attachment to a specific place" and "sense of community" help our understanding of how individuals and groups create bonds to a specific place. Although the "size" and the comfort of the habitable space are essential for residential satisfaction, other aspects of the living conditions also modulate its importance. Social integration in the neighborhood, the availability of satisfactory services, the existence of green space, the aesthetics of the built environment, and easy access to the public transportation network encourage the affective investment of the neighborhood and enhance people–environment congruity.

The more this congruity is important, the more the neighborhood matches one's aspirations and allows one to satisfy his or her material and social needs, and the more one develops attachment toward one's neighborhood. In certain cases, the interpersonal relationships that the individual develops within his or her place of living are more important for the feeling of belonging and attachment than the physical specifications of the urban environment. This is for instance the case in shantytowns, in which forced dislocations and allocations of far more comfortable new dwellings are regarded as traumatic by the inhabitants.

The feeling of being "at home" is closely connected to satisfaction with one's neighborhood. The degree of satisfaction with the environmental attributes of one's neighborhood (i.e., the green spaces, the aesthetics of the built framework and the degree of noise) determines both the intensity of the affectivity that one develops toward the neighborhood and one's feeling of well-being.

For individuals who have already acquired "basic" living conditions and who have an income that allows them to achieve a good quality of life, the agreeable character of the neighborhood has a modulating effect on satisfaction concerning available space in the dwelling. The feeling of being at home in one's neighborhood is linked to the frequency of encounters with others in the neighborhood, the number of close relations, the nature of local relationships, and the satisfaction with them. This explains why social relations related to a determined place may be more important to feelings of attachment than the place itself.

Besides the home and neighborhood environments, other urban environments involve congruity between people and their environment, including workplaces and institutional environments such as hospitals, prisons, and homes for children or the elderly. How can these environments be designed to meet the needs of their occupants?

5.2. Control of Density in a Collective Habitat

High density is an environmental characteristic often regarded as responsible for the individual's withdrawal, inattention to others, and aggressive behaviors. The most specific interventions concern the control of density and the regulation of behavior in spaces open to the public.

Certain interventions involving the design of space are effective in diminishing the feeling of being constrained, notably in collective buildings, prisons, public libraries, and cafeterias. Some interventions may explicitly refer to the principles elaborated by Barker in his ecological theory when organizing public spaces or spaces open to the public.

The larger the collective buildings, the less social activity the residents tend to have (neighborhood relations, mutual help, and associative involvement). This is usually attributed to the management of intimacy in public spaces. For instance, inhabitants of buildings with long corridors feel threatened with regard to intimacy. An architectural intervention consisting of fractioning long corridors by introducing doors diminishes the feeling of crowding, enhances the feeling of control, and favors interactions in the spaces so delimited.

5.3. Management of Neighborhood Spaces

Contributions to the management of neighborhood spaces involve the "habitability" of neighborhood (districts) and the treatment of environmental problems and insecurity. Such interventions mainly focus on neighborhood revitalization and interventions designed to improve residents' satisfaction and reduce crime and feelings of insecurity.

5.3.1. Neighborhood Revitalization

Neighborhood revitalization has to follow two apparently contradictory principles: (i) to foster attachment and a feeling of belonging to a well-delimited and clearly identifiable area and (ii) to open the area in order to connect people to the city.

The key to successfully revitalizing certain districts is in the way in which these two contradictory exigencies are handled. Such interventions have to be conducted jointly with regard to both the social and the physical aspects of the environment and engage all the relevant actors (i.e., allow cooperation between the inhabitants, the sponsors of the project, and the municipality).

Interventions concerned with the development of urban residential environments consist of improving sidewalks or pathways for inhabitants. Their aims are (i) to encourage pedestrians to use proximate pathways in order to promote social interaction among inhabitants and foster control of different outdoor activities, (ii) to diversify the outdoor environment in order to

increase walking in the whole proximate environment, (iii) to give residents the opportunity to engage in different activities (shopping, recreational opportunities, walking, etc.), and (iv) to reduce the number of cars passing through the area and limit their speed ("quiet zones"). Such interventions often require redesigning the district as well as the built environment so that dwellings face toward roads and the installation of green spaces with multiple recreation opportunities for all residents.

Similarly, the concept of defensible space, introduced in 1972 by Newman, has often been used as a guideline for the design of proximate environments. Development designed to give maximum control over public space to inhabitants, by erecting symbolic or real barriers, restricting access to the district, designing well-illuminated pathways, and encouraging the individual appropriation of dwellings by differentiating them from each other, has given positive results. Postrenovation evaluations have found less maintenance costs and less reprehensible behaviors.

Such interventions correspond to a territorial logic. Encouraging outdoor activities and restricting vehicles allow a neighborhood to define its identity and to appropriate and therefore better control its immediate environment and the street, which enhance the feeling of belonging and increase the feeling of being safe.

5.3.2. Environmental Annoyances

Dealing with environmental annoyances, such as excessive noise or undesired odors, often requires an intervention at the source and socioenvironmental management based on an intervention to improve all aspects of the neighborhood so that residents' well-being is improved and crystallization on the incriminated annoyance is avoided.

The same type of intervention is necessary when a new industry is planned, especially if it is reputed to create new and specific annoyances. In this case, particular attention must be paid to the possible NIMBY (not in my back yard) phenomenon (i.e., preliminary rejection by inhabitants).

5.3.3. Insecurity and the Feeling of Insecurity

The feeling of insecurity and subsequent fear are seldom correlated with objective dangers and are essentially based on visual cues. Fear of crime is exacerbated by an urban environment that appears to be poorly maintained, characterized by littering, vandalism and graffiti, uncleanness, lack of illumination, and deviant behavior. Cleaning up vandalized buildings and urban furniture reduces littering, and other interventions focused on the signs of disorder and lack of control in the neighborhood provide an increased feeling of security.

A pleasant environment may also contribute to the feeling of security of inhabitants. Nasar repeatedly showed that coherence, compatibility, congruity, legibility, and clarity of the urban environment enhance the evaluative quality of cities, downtown street scenes, and residential scenes. On the other hand, a design that obviously impedes a certain behavior in the sense of defensible space may be interpreted as a sign of danger and produce a feeling of insecurity.

6. URBAN PUBLIC ENVIRONMENTS: THE CITY

Regarding public urban space, environmental psychology may intervene in two ways: (i) management of the urban structure by improving the legibility of the urban environment and revitalizing the urban public space and (ii) management of the urban social diversity. The main aim of both types of intervention is to favor appropriation.

6.1. Management of Public Urban Spaces

The management of urban environments mainly consists of encouraging connections between urban districts and giving urbanites the opportunity to appropriate the various public spaces for a large diversity of activities, including recreation and contact with nature.

6.1.1. Homogeneity and Legibility of the Urban Space

The urban structure is not homogeneous. Beyond the classical distinction between the center and the suburbs, several zones may be identified. One can distinguish fluid zones, to which individuals do not anchor because such zones are only designed for mobility; anchorable zones (i.e., living places) with which people can identify; and functional zones, which are similar to fluid zones (people simply "go through" them).

Furthermore, the natural fractures (e.g., hills and rivers) or artificial ones (e.g., highways and rail tracks) may shape the urban landscape and isolate districts from one another both physically and psychologically.

According to Lynch's principles of urban structure, the mental map methodology allows the problems of isolation to be diagnosed and the means of redesigning urban layouts of districts or zones to be proposed in order to improve their interconnectedness.

Legibility refers to another, mostly neglected aspect of the environment—that is, the way collective memory and cultural references shape the perception of the urban environment and, consequently, enhance people's appropriation. In identifying "landmarks" such as monuments or geographical particularities, mental maps contribute to the understanding of the legibility of the city.

6.1.2. Parks, Squares, and Nature in Urban Environments

The gentrification of public spaces concerns not only the design of sidewalks and streets in order to encourage their use by pedestrians but also the interventions on parks and squares. Urban squares often suffer either monospecific or a deficit of frequentation due to the limited activities that they offer. Providing an increased number of activities and introducing water (fountains), trees for shade, and benches for resting and interaction make them attractive for passer-byes and opens them to a variety of different occasional and regular users.

Urban residents often seek nature and want to visit urban parks, gardens, and recreational areas for leisure. Green spaces and the natural environment can provide not only an aesthetically pleasing setting but also restorative experiences, including a positive effect on health. Gifford identified the following main benefits of nature: cognitive freedom, escape, the experience of nature, ecosystem connectedness, growth, challenge, guidance, sociability, health, and self-control. The most important seem to be the sense of freedom and control felt in nature, in contrast to an urban environment, which is often perceived as constraining.

Behavioral zoning, which consists of allocating specific spaces to specific activities, allows access to be controlled and canalizes the flux of visitors, thus avoiding conflicts resulting from different aspirations and contradictory needs (rest/exercise) and reducing the feeling of crowding. Determining the mean distances from habitat to where people perform different leisure activities is important in the planning of neighborhood leisure places.

Concerning children, research has consistently found a lack or absence of places for them to frequent. Gaster proposes guidelines for the planning of proximal spaces for children.

6.2. Management of Urban Diversity

Individuals often group spontaneously or are grouped within the urban territory according to social, cultural, ethnic, and economic affinities and form more or less homogeneous districts that tend to singularize within the urban community in its totality. They often express a feeling of belonging and existing as a specific group through urban territorial identification.

Socioenvironmental management of urban diversity means not only encouraging the coexistence of different cultures and fostering harmonious development of all minorities but also ensuring an equitable and just repartition of services in order to provide the different communities with the infrastructures they need. Only cooperation between inhabitants and authorities as well as participation in decision-making encourage the feeling of belonging to a distinct urban community. This is achieved if two aspects are promoted: the public expression of difference and its acceptance and the construction of a common destiny through the feeling of belonging to the urban community. These objectives require intervention with regard to four aspects. First, efficient infrastructure and equitable access to different urban services (health, transport, energy, water, garbage collection, and so on) must be provided, and these services must be recognized as being a common good contributing to the well-being of everyone. Second, the needs, expectations, and aspirations of the different components of the urban population must be identified, regular contact with the different communities (and/or their representatives) of the metropolitan population must be maintained, and the different communities must be informed of any new development and/or improvement of services and infrastructures. Third, a regular and easily accessible system of communication between the town administration and the population must be developed. Finally, the expression of cultural particularities through the educational system, the media, public and cultural activities must be taken into account and encouraged, and mutual recognition among the different communities must be encouraged by promoting intercultural sports or artistic encounters.

Such principles gain significance only if they are articulated together in order to allow the simultaneous

expression of differences and a feeling of belonging to an unique urban community that has cultural identity, on the one hand, and is transversal and metropolitan, on the other hand. Such interventions encourage not only local appropriation and identification but also the sense of belonging to an identified urban community.

Regarding the city as a place of living, creating the conditions of its appropriation leads to identity building in the sense of urban citizenship, which meets the requirements of sustainability insofar as this is, according to the Cities, Identity, and Sustainability model proposed by Pol, a condition of people's engaging more readily in pro-environmental behavior.

7. TOWARD SUSTAINABILITY OF URBAN ENVIRONMENTS

Sustainable development and globalization are likely to have a major effect on the development of urban settlements. Both are key concepts in the economics of the 21st century, inevitably having an impact on our ways of living in general.

7.1. Globalization and Sustainability

By defining sustainable development as development capable of satisfying the needs of the present generation without compromising the ability of future generations to satisfy their own needs, the Bruntland Report, published in 1987, opened the way to questions related to quality of life. The reference to needs allows not only the requirement for development to be harmonious and respectful to the environment but also that the individual's well-being be recognized. People-centered analyses focusing on well-being raise the crucial questions of attachment, control, and identity or identities.

Globalization and its corollaries, global trade and communication, create pressure toward cultural uniformity in lifestyles. Globalization and the anxieties that it engenders also require the local or regional priorities and particularities, cultural differences, and therefore also specific needs to be recognized. Cultural differences and identity are essential issues with regard to the relationship between people and their environment. Identities give structure and meaning to life. They are at the origin of both conflict and cohesion, and they strongly affect patterns of both exclusion and solidarity. At both the interindividual and the societal level, gender, religious, and cultural identities generate new forms of solidarity, often on a transnational or transcultural level. Identities are also frequently related to the environment. They shape relationships between people and environment at different spatial levels, depending on appropriation and spatial anchoring.

7.2. Sustainable Cities

The requirement for cities to be sustainable raises further challenges. Cities host increasingly culturally heterogeneous populations. The implementation of sustainable development through Local Agenda 21 initiatives will only be possible if the local communities reach a consensus. The current trend aims for proactive consensus-building approaches that attempt to involve people in the decision-making process. Meanwhile, many Third World communities are involved in self-help initiatives, but they are rarely able to participate in wider urban planning or politics. Does the concept of citizenship explicitly impose a single, egalitarian political identity on the disparate and separate identities likely to exist within a population of any significant size? How may the intracommunity differences be recognized and how may they converge in a common urban citizenship (Fig. 2)?

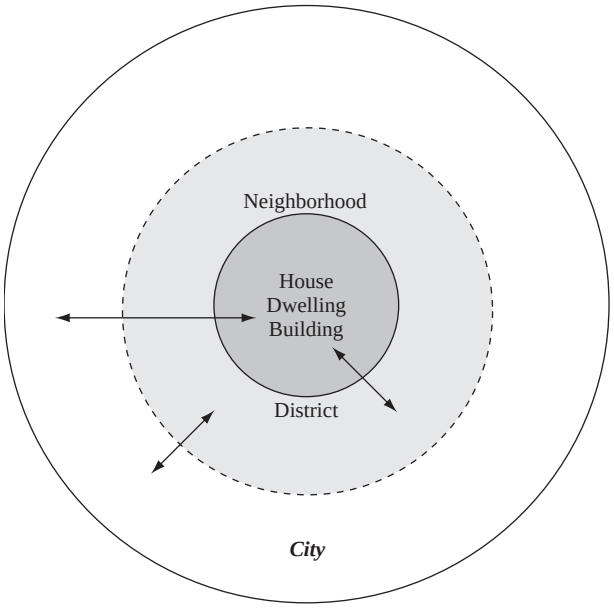


FIGURE 2 From dwelling to citizenship: extensions of appropriation.

In other words, the challenge for environmental psychology is to provide tools not only able to foster the appropriation at the dwelling or the neighborhood level but also to encourage the appropriation of the city and help the individual to identify with the urban community as a whole. Placing the citizen at the center of public life is an important challenge. Achieving sustainable cities requires strong civic cultures and a new politics of cohesion and cooperation.

See Also the Following Articles

Environmental Design and Planning, Public Participation in ■ Environmental Stress ■ Person–Environment Fit ■ Privacy ■ Residential Satisfaction and Perceived Urban Quality ■ Territoriality

Further Reading

Altman, I. (1975). *The environment and social behavior*. Monterey, CA: Brooks/Cole.

Altman, I., & Low, S. M. (1992). *Place attachment. Human behavior and environment: Advances in theory and research* (Vol. 12). New York: Plenum.

Barker, R. G. (1968). *Ecological psychology: Concepts and methods for studying the environment of human behavior*. Stanford, CA: Stanford University Press.

Canter, D. (1977). *The psychology of place*. London: Architectural Press.

Gifford, R. (1997). *Environmental psychology: Principles and practice* (2nd ed.). London: Allyn and Bacon.

Kaplan, S. (1995). The restorative benefits of nature—Toward an integrative framework. *Journal of Environmental Psychology*, *15*(3), 169–182.

Krupat, E. (1985). *People in cities: The urban environment and its effects*. Cambridge, MA: Cambridge University Press.

Low, S. M., & Altman, I. (1992). *Place attachment*. New York: Plenum.

Milgram, S. (1970). The experience of living in cities: A psychological analysis. In F. F. Korten, S. W. Cook, & J. I. Lacey (Eds.), *Psychology and the problems of society*. Washington, DC: American Psychological Association.

Moser, G. (1992). *Les stress urbains*. Paris: Colin.

Moser, G., Pol, E., Bernard, Y., Bonnes, M., Corraliza, J., & Giuliani, V. (2002). *People, places & sustainability*. Göttingen, Germany: Hogrefe & Huber.

Moser, G., & Uzzell, D. (2003). Environmental psychology. In T. Millon, & M. J. Lerner (Eds.), *Comprehensive handbook of psychology. Vol. 5: Personality and social psychology* (pp. 419–445). New York: Wiley.

Moser, G., & Weiss, K. (2003). *Espaces de vie: Aspects de la relation homme–environnement*. Paris: Colin.

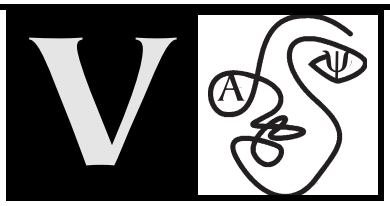
Preiser, W., & Ostroff, E. (2001). *Universal design handbook*. New York: McGraw-Hill.

Proshansky, H. M., Ittelson, W. H., & Rivlin, L. G. (1970). *Environmental psychology: Man and his physical setting*. New York: Holt, Rinehart and Winston.

Simmel, G. (1957). The metropolis and mental life. In K. H. Wolff (Ed. and Trans.), *The sociology of Georges Simmel*. Glencoe, IL: Free Press.

Stokols, D., & Altman, I. (1987). *Handbook of environmental psychology*. New York: Wiley.

Wicker, A. W. (1979). *An introduction to ecological psychology*. Monterey, CA: Brooks/Cole.



Values and Culture

Peter B. Smith
University of Sussex, Brighton, United Kingdom

1. Introduction
2. Levels of Analysis
3. Individual-Level Dimensions of Values
4. Culture-Level Dimensions of Values
5. Conclusion
Further Reading

GLOSSARY

country-level analysis Studies in which values data for each nation are aggregated so that each country's data become a single data point.

ecological fallacy The belief that because two variables are correlated in a particular way at one level of analysis, they must show the same relationship to each other at another level of analysis.

individual-level analysis Studies in which variations in values between individuals within a population are examined.

response bias A tendency to consistently use the extremes (or the midpoint) of scales measuring agreement or disagreement with questionnaire items, regardless of item content.

value dimension A continuum summarizing the relative endorsement of values or value types that contrasts some value types with other value types.

values Abstract generalized beliefs about desirable goals and about ways of achieving those goals.

value type A group of individual values shown empirically to cluster together.

Comparisons of values have provided one of the most useful ways of understanding cultural differences. This is because there is less variation between different cultures in the meaning of particular values than there is in the meaning of specific behaviors. When individuals are studied, there is evidence of universal linkages between particular values and certain self-concepts, attitudes, and behaviors. However, within different nations the array of those holding particular values is not the same. By studying culture-level value differences, the context that influences individuals within each culture can be better understood.

1. INTRODUCTION

The study of values has steadily become more central to cross-cultural psychology as it has developed over the past several decades. This is largely a consequence of the way in which many researchers have chosen to conceptualize cultural differences. Early researchers placed emphasis on identifying differing behavior patterns and treating these as indicative of cultural difference. However, no generalizable patterns of behavioral difference were identified. Recently, it has been noted that similar behaviors may be interpreted in quite different ways within different cultural groups. Thus, cultural difference may depend more directly on the ways in which behaviors are interpreted, evaluated, and given meaning than on their simple presence, absence, or

©2004 Elsevier Inc.
All rights reserved.

frequency. If this perspective is adopted, then the study of values becomes of central importance since a person's values comprise an abstract generalized template, in terms of which all events, experiences, and possibilities are likely to be evaluated.

Individuals may endorse many values and their actions in any particular situation may well depend on the relative strength and relevance of each value to that situation. Thus, if values are to provide a secure theoretical basis for cross-cultural psychologists, they must be surveyed comprehensively, rather than in terms of some specific value that achieves a fashionable status among researchers.

Values are typically measured by providing respondents with a list of single words or phrases representing abstract values, such as honesty, authority, kindness, and so forth, and asking how strongly each value is endorsed, using rating scales. This procedure follows the pioneering work by Rokeach in the United States, who asked respondents to rank a list of values in order of their importance. However, some of the researchers whose work is surveyed in this article relied in part on respondents' statements evaluating specific behaviors and expressing attitudes toward hypothetical events, from which they inferred the presence of particular underlying values.

Before measures of values can be used to understand cultural difference, there are two key methodological problems that must be resolved. First, it is necessary to take account of the different ways in which respondents in different cultures respond to questionnaires. Members of some cultural groups show different types of response bias than do others. For instance, Hispanic respondents give more extreme ratings than Caucasian Americans. There may also be differing norms as to whether one should present oneself modestly, as in Japan, or assertively, as in the United States. Several procedures for correcting this type of bias have been proposed. The most widely used are centering, in which each score is subtracted from the respondent's overall mean, and within-subject standardization, in which centered means are also divided by the standard deviation. Some critics argue that these types of standardization actually eliminate variance that is substantive rather than artifactual. An alternative strategy is test hypotheses both with and without standardization. If hypotheses are more strongly supported after standardization, the argument is strengthened that bias was artifactual.

The second problem to be overcome is to ensure that the particular values included in a questionnaire are interpreted in equivalent ways within each of the cultures being sampled. This requirement has often been neglected. Equivalence can be tested by determining whether endorsement of the various values is intercorrelated in a similar way within each of the cultures. For instance, if endorsements of "love," "friendship," and "honor" show a similar pattern of relationships in each culture, the chances are increased that these concepts have equivalent meanings across cultures.

2. LEVELS OF ANALYSIS

Most psychological studies concern the behavior of individuals. Values are also measured by seeking responses from samples of individuals. It may seem paradoxical that cross-cultural psychologists should have found it useful to survey values in order to understand cultural differences, when these differences are not simply between individuals but between large aggregates of individuals. The distinction between culture-level analysis and individual-level analysis is crucial to cross-cultural psychologists, and it is frequently not well understood. Hofstede was among the first to note its importance. In identifying the four value dimensions along which he classified national cultures, he used data aggregated to the culture-level, not individuals' responses. Thus, each nation's data became one case, not many. Hofstede maintains that his dimensions are applicable only to the interpretation of cultural phenomena, and not to the behavior of individuals; however, many others have used his concepts for individual-level analyses.

There is thus a prevailing cultural milieu in which each individual resides. Any one individual may have values that are congruent with that milieu or that diverge from it. Furthermore, two values that can be found together within a cultural milieu may be such that any one individual could not endorse both. For instance, a culture may have prevailing values favoring both authority and humility, but any one individual will have to either seek authority or endorse humility. To assume that variables that are positively correlated at one level of analysis will also be similarly correlated at the other level is to commit the so-called ecological fallacy.

The relations of culture-level values to other variables may also differ from those found at the individual level. For instance, scores on Hofstede's culture-level individualism variable correlate positively across 45 nations with mean scores for job satisfaction. In contrast, individual-level analysis in Hong Kong revealed a negative correlation between high job satisfaction and individualist values. Therefore, when surveying studies of values across cultures, levels of analysis must be distinguished.

3. INDIVIDUAL-LEVEL DIMENSIONS OF VALUES

The principal attempt to identify individual-level dimensions of values that are replicable cross-culturally is that by Schwartz, whose approach derives from that pioneered by Rokeach. Schwartz grouped 56 single values into value types according to their common goals. He reasoned that the basic human values likely to be found in all cultures are those that represent universal requirements of human existence (biological needs, requisites of coordinated social interaction, and demands of group functioning) as conscious goals. Drawing on the values identified by previous researchers and writers, he grouped them into the following motivationally distinct value types:

1. Power—social status, dominance over people and resources
2. Achievement—personal success according to social standards
3. Hedonism—pleasure or sensuous gratification
4. Stimulation—excitement and novelty
5. Self-direction—independence of thought and action
6. Universalism—understanding, tolerance, and protection for the welfare of all people and nature
7. Benevolence—preserving and enhancing the welfare of people to whom one is close
8. Tradition—respect and commitment to cultural or religious customs and ideas
9. Conformity—restraint of actions and impulses that may harm others and violate social expectations
10. Security—safety and stability of society, relationships, and self

Schwartz then postulated that actions taken in pursuit of each value type have consequences that may conflict or may be compatible with the pursuit of other types. For example, pursuing achievement values may conflict with pursuing benevolence values: Seeking personal success may prevent enhancing the welfare of others. The theoretical pattern of conflict and compatibility among value types is portrayed in Fig. 1. Competing value types are placed opposite one another, with more complementary types in closer proximity.

Schwartz confirmed this structure empirically, revealing that the 10 value types can be summarized as the two bipolar dimensions also shown in Fig. 1: Openness to Change (including self-direction and stimulation) versus Conservation (security, conformity, and tradition) and Self-Enhancement (power and achievement) versus Self-Transcendence (universalism and benevolence). Hedonism includes elements of both Openness to Change and Self-Enhancement.

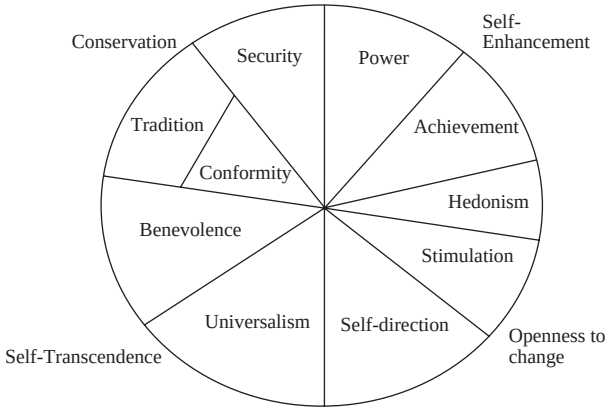


FIGURE 1 Schwartz's value types.

Schwartz's data were collected using a survey comprising 56 values selected to represent each of the value types. The Schwartz Value Survey (SVS) has been administered in more than 60 nations, sampling students and urban schoolteachers. Respondents rate the importance of values as guiding principles in their lives. Correspondence between theorized and actual structure of value types was assessed by comparing two-dimensional spatial representations of the intercorrelations among the values to the theoretical spatial arrangement in Fig. 1. Separate analyses have been reported for more than 100 samples. These analyses support the near universality of the 10 value types and of their structural relations. The two bipolar dimensions are found in almost every sample., The Portrait Values Questionnaire comprises verbal portraits of persons exemplifying the 10 value types. The structure of values found earlier has proved replicable using this measure within samples from Israel, Uganda, Italy, and South Africa. In addition to the uniformity found in the structure of values, the SVS data set also reveals that certain value types are universally more strongly endorsed than others. In particular, benevolence and self-direction are favored over power and tradition.

Further analyses of SVS data have indicated some specific variations in the meaning of individual values

between nations. For instance, in Japan "true friendship" correlated strongly with security values, whereas in most other samples it correlated better with benevolence values. Australians had a rather different conception of loyalty and responsibility than respondents from other nations. In Spain, "honor" was associated with family and social interdependence, whereas in The Netherlands it was associated with self-achievement and autonomy. These variations in the overall results are valuable in identifying what may be distinctive about particular cultures or regions. However, the principal benefit from the individual-level Schwartz project is that it provides a valid measure with which to test whether values have similar correlates in different cultures.

3.1. Correlates of Individual-Level Values

Young people in a wide variety of nations attribute greater importance to Openness to Change and less importance to Conservation than their elders. More educated people show the same value patterns as younger people. A meta-analysis of samples from 47 countries found that women more strongly favor security and benevolence values, whereas men favor self-direction, stimulation, hedonism, achievement, and power values. Differences are strongest for the benevolence and power value types. However, the magnitude of gender differences is much greater in some cultures than in others. The Chinese Value Survey data from 23 nations also found numerous culture by gender interactions. Even the largest gender differences are smaller than differences in values typically found in relation to age, education, and national culture.

3.2. Attitudes and Behavior

If individuals' values can be shown to consistently predict attitudes and behaviors in a variety of different cultures, then the case is strengthened for using variations in individual values to interpret cultural differences. If the relations between individual values and other phenomena are found to vary between cultures, then other ways should be sought to explain differences between cultures. For instance, it might be found that acting consistently with one's personal values is esteemed in some cultures, whereas conforming to social expectations is preferred in others. A country-level meta-analysis of Asch conformity studies conducted in 17 nations did find such variations. These variations might enhance or weaken the values–behavior link. At the individual level, values have been compared with the Crowne–Marlowe measure of need for social approval in Finland and Israel. Despite the differences between these nations, need for social approval was consistently positively correlated with conformity values and negatively with hedonism and stimulation.

Religious commitment has been studied among Muslims, Jews, and Catholic, Orthodox, and Protestant Christians in 13 countries. Religiosity is consistently correlated most positively with tradition values and most negatively with hedonism and stimulation. The correlations are stronger in Western Europe than in the former communist states of Central and Eastern Europe. Religiosity and universalism values are more negatively correlated in the Orthodox and Muslim samples than elsewhere. Thus, there is evidence for consistent linkages between values and religiosity, but also some effects attributable to variations in church—state relations.

Consistency has also been found between SVS values and support for particular political parties. In Hungary, Israel, The Netherlands, and Spain, the values in opposing positions in the SVS predicted support for opposing political parties. In the United Kingdom, Japan, New Zealand, and the United States, universalism values correlated positively with anti-nuclear activism, whereas power values correlated negatively. Correlations for self-direction with universalism were mostly positive and those for security and conformity mostly negative. These results indicate the importance of universalism and self-direction versus power, security, and conformity for predicting political attitudes and behavior.

SVS scores have been compared with a self-concept measure among students in Singapore and the United States. Consistent links were found between power values and vertical individualism, self-direction values and horizontal individualism, conformity values and vertical collectivism, and benevolence values and horizontal collectivism.

Preferred styles of managing conflict among management students in India, the Philippines, Hong Kong, and the United States suggested that respondents' individual-level Conservation scores mediated their preference for avoiding, whereas Self-Enhancement and Openness to Change mediated preference for competing.

The studies reviewed in this section indicate that there is a growing body of evidence of cross-national consistency between endorsement of individual values and a variety of demographic indicators, self-concept, attitudes, and social behaviors. This enhances the possibility of explaining some cultural differences in terms of the differing spectrum of values endorsed by individuals from different areas of the world. However, such explanations will always be no more than partial since they ignore the likely effects of culture-level variables.

4. CULTURE-LEVEL DIMENSIONS OF VALUES

The average values that are found within a particular nation, or within discrete cultural groupings that comprise a nation, define a context within which individuals are born and socialized. No national culture is entirely homogeneous, and contemporary nations are characterized by a rich variety of ethnic, educational, occupational, professional, and regional subcultures. Consequently, it is normal rather than exceptional that there should be divergence between individuals' values and the average values found to characterize those individuals' nations as a whole. Although psychologists have mostly preferred explanations of human behavior in terms of variables derived from the person, attention to culture-level variables opens up some of the more neglected situational determinants of behavior. The values espoused by most of those around us will mould and temper the ways in which we act on our own values.

It might be argued that nations are products of history and politics, and that many nations would fail to satisfy a definition of culture that uses shared evaluations of behavior as a defining criterion. However, modern nations are characterized by numerous mechanisms by which they seek to create and maintain cultural homogeneity, including education, legislation, language policies, and the mass media. Almost all researchers of culture-level phenomena have equated cultures with national cultures.

There have been numerous attempts to characterize national cultures. The current discussion highlights those that have focused most directly on values.

The pioneer study was Hofstede's well-known study of IBM employees from 50 nations and three regions; Hofstede provided an updated account in 2001. Of the four culture-level dimensions identified by Hofstede, individualism–collectivism has proved by far the most influential. This article gives greater attention to the three remaining dimensions. The four dimensions were determined by aggregating responses within each nation to each questionnaire item and then factor analyzing the resulting country scores.

Since the original questionnaire was not designed on the basis of any explicit theory, it is perhaps not surprising that the resulting factors comprised responses to a mixture of items. High scores on Power Distance indicated, for instance, that respondents reportedly feared to express disagreement with their manager, perceived their boss as autocratic or paternalistic, and did not want a boss who is consultative. High scores on Individualism versus Collectivism were defined by emphases on work goals of freedom to work in one's own way, sufficient time for personal life, and challenging work versus work goals of good working conditions, opportunities for training, and full use of skills and abilities. High scores on Uncertainty Avoidance reflected high reported job stress, beliefs that company rules should not be broken, and expectations of continuing to work for IBM for a long time. High scores on Masculinity versus Femininity reflected emphases on earnings, recognition, advancement, and challenge in one's work versus good relations with the boss and others, living in a desirable area, and having high job security. The items defining Individualism–Collectivism and Masculinity–Femininity are the only ones that can be considered as direct measures of values, but Hofstede asserts that the other dimensions do reflect underlying values. He has sought to refocus attention on the more neglected of his dimensions by the publication of a substantially revised edition of his 1980 book, which gives equal attention to all dimensions. In the new edition, his arguments are supported by citation of the large number of studies relevant to each dimension that have been published during the past two decades. He compares country scores on a variety of indices with value scores for the relevant nations. However, few of these studies have actually attempted to replicate the Hofstede measurements. The validity of all such comparisons rests on the validity of measurement in the original study and on the degree to which those measures continue to accurately represent particular nations.

Michael Bond was the first to test the replicability of Hofstede's dimensions. He constructed a Chinese Value Survey (CVS) intentionally based on Eastern rather than Western values by asking Chinese scholars to nominate values "of fundamental importance in Chinese culture."

Sampling students from 23 countries, a country-level factor analysis yielded four orthogonal factors, three of which correlated well with the Hofstede scores for the 20 overlapping nations. No CVS factor correlated with Uncertainty Avoidance. Bond called his fourth, unique CVS factor Confucian Work Dynamism, and Hofstede adopted this as an added value dimension of cultural variation missing from his theory. Hofstede interpreted it as opposing a dynamic orientation toward the future to a more static orientation toward the past and present and therefore renamed it Long-Term Orientation. He suggests that Western cultures are preoccupied with the search for Truth so that the emergence of an Uncertainty Avoidance factor from a survey designed by Western psychologists is to be expected. In contrast, he suggests that Pacific Asian cultures are preoccupied with the search for Virtue, which might explain a greater emphasis on Long-Term Orientation or Confucian Work Dynamism among the values identified by Chinese scholars.

Schwartz's individual-level analyses have already been described. By aggregating his individual-level data, he has also identified culture-level variations in values. Only those 45 values demonstrated to have relatively stable cross-cultural meanings in studies within each nation were included. Multidimensional scaling was again used. At the culture level, he argues that the content of cultural value dimensions reflects alternative solutions to three basic societal issues:

1. Relations between individual and group: To what extent are persons autonomous versus embedded in groups?
2. Ensuring responsible social behavior: How can people be motivated to consider others' welfare and coordinate with them?
3. The role of humankind in the natural and social world: Is it more to submit, to fit in, or to exploit?

The results consistently yielded seven value types. These can be summarized as three bipolar dimensions, reflecting the three issues detailed previously: Conservatism versus Autonomy (divided between intellectual autonomy and affective autonomy), Hierarchy versus Egalitarianism, and Mastery versus Harmony. Nation scores were computed by combining the mean ratings on the values designated a priori as representing each value type. Scores were also computed for each dimension (one polar value type minus the opposing value type). The scores derived from teacher samples correlated well with those obtained from student samples across 40 nations.

A further survey was conducted by Trompenaars, who sampled the work-related values and attitudes of approximately 15,000 organization employees from more than 50 countries. Many of his questionnaire items were based on the earlier theories of the U.S. sociologist, Talcott Parsons. Trompenaars provided country means for responses to individual questionnaire items and suggested several cultural value dimensions, but he undertook no empirical analyses to assess the validity of these dimensions. Smith, Dugan, and Trompenaars selected the items from his questionnaire that were suitable for the empirical analysis of dimensions. These items measured preferences for using universalistic versus particularistic standards in relationships, for granting status on the basis of ascribed versus achieved qualities, for orientation to the interests of the self versus the collective, and for different styles of organizational management. After standardization of individual responses, item means for 43 nations were entered into a multidimensional scaling analysis. Two clear dimensions emerged: Conservatism versus Egalitarian Commitment and Loyal Involvement versus Utilitarian Involvement.

Inglehart has also identified two culture-level value dimensions based on factor analysis of aggregated data from public opinion surveys conducted within 43 nations. The first, Well-Being versus Survival, shows a strong positive correlation with Hofstede's individualism scores and strong negative correlations with Power Distance scores. The second dimension, Secular–Rational versus Traditional Authority, shows similar but weaker correlations.

4.1. Relations among Culture-Level Dimensions

There are five studies that have each sampled values from at least 23 nations. These studies vary in methodological strength. Only Schwartz derived his items from both Western and non-Western sources, controlled for meaning equivalence, and demonstrated replicability of his dimensions across samples. Nonetheless, consistencies across the different studies may point to the most central and reliable dimensions of cultural values. Smith and Schwartz examined the correlations between those country scores that were common to each pair of surveys. They conclude that there is substantial convergence between results, even though different authors use different names for

dimensions of variance in values. The correlations point to a close positive association between two basic dimensions identified in different ways by different authors:

1. The preferred cultural view of individual–group relations (autonomous versus embedded)
2. The preferred cultural mode of motivating responsible social behavior and allocating resources (negotiation among equals versus acceptance of unequal, hierarchical roles).

These dimensions are virtually identical to the first two culture-level dimensions identified by Schwartz. However, it is unlikely that all the ways in which values vary across cultures have been identified.

4.2. Studies Related to Culture-Level Value Dimensions

As noted previously, categorizing cultures in terms of nations involves substantial arbitrariness. Within samples from large numbers of nations, it is possible that discernible clusters of nations can be identified, within which similar values prevail. Hofstede reports a hierarchical cluster analysis of his data set of 53 nations and regions. He lists 12 "culture areas," but some of these are less clearly meaningful than the 7 areas identified in his 1980 book, which included only the 40 nations with large samples: Nordic (e.g., Norway and Sweden), Anglo (e.g., the United States, the United Kingdom, and Australia), Germanic, Near Eastern, less developed Asian, less developed Latin, and more developed Latin. Japan stood alone as a more developed Asian nation. Schwartz conducted multidimensional scaling analyses using nation scores on his seven culture-level value types using data from both his teacher and his student samples. Six relatively similar culture areas emerged from both analyses: West European, Anglo, East European, Islamic (e.g., Turkey and Indonesia), East Asian, and Latin American. Japan was again unique. The similarities between the Hofstede and Schwartz culture clusters, and some of the differences, simply reflect the sampling, such as the lack of former Soviet bloc nations in the IBM study and underrepresentation of Islamic countries in both. Other differences may reflect cultural change during the past 20 years. One study including 43 countries found two clear culture areas, one for Central and Eastern Europe and the other for northern Europe. Somewhat less clear Far Eastern, Latin, and Anglo areas were also discernible. The convergence of these results with those reported by Hofstede and Schwartz is substantial.

We can therefore have some confidence in how to characterize the value preferences of regional groupings throughout the world. It is less clear how great can be the contribution of psychologists to explaining why these differences are as they are. The principal tools that can be employed are correlations and regressions, and at the country level these are most useful when large numbers of nations have been sampled. Furthermore, inferring causation from any relationships found is notoriously unreliable. To take a widely cited example, country gross national product correlates +.85 with Hofstede's Individualism and –.65 with his Power Distance, but these figures do not indicate whether wealth is a cause or a consequence of such values or, indeed, whether the high correlations are a side effect of some unspecified third factor. We know that economic growth during the latter part of the 20th century was correlated +.70 with scores for Confucian Work Dynamism, but this correlation does not hold for the most recent decade, maybe because political and economic factors are now different.

A statistical technique (regression equations) permits the prediction of a value profile from other variables, such as attributes of the countries that have been studied. This technique has been used to test explanations of value profiles in Western and in Eastern Europe. Hofstede computed regressions that summarized what he called system-level indicators of each of his value dimensions. Moderate latitude, high population, and low wealth explain 56% of variance in countries' Power Distance scores. Low wealth explains 8% of variance in Uncertainty Avoidance. High wealth and extreme latitude explain 78% of variance in Individualism. High population explains 8% of variance in Masculinity scores.

4.3. Correlates of Culture-Level Values

An alternative procedure is to examine data on the extent of a specific psychological phenomenon across nations and determine whether country-level value scores can account for the variance that has been found. This is a hazardous enterprise since the country scores are derived from samples that differ from those whose behavior is being studied. A statistical technique (meta-analysis) permits the integration of data from

many studies. It has shown that there is greater conformity in countries that score high on Hofstede's Collectivism and low on Schwartz's Intellectual and Affective Autonomy. However, it could be that the behaviors identified as conformity in cultures high on Individualism and Autonomy may be thought of as sensitivity to others in cultures high on Collectivism and Conservatism.

Culture-level values have also been linked to work-related variables. Among managers in 21 nations, role overload is predicted by high Power Distance, and role ambiguity is predicted by low Power Distance. Schwartz found work to be more central to people's lives where Hierarchy and Mastery score high. He also found that the tendency to view work as an obligation is greater in cultures high on Conservatism and Hierarchy. Across 47 nations, it has been found that the sources of guidance on which managers relied most in handling a range of work events could be predicted from country scores derived from Hofstede, Trompenaars, and Schwartz. Hofstede cites numerous additional significant correlations between his dimensions and data derived from opinion surveys, objective indices, and psychological studies. Some of these studies also sampled substantial numbers of nations. It is clear that at the culture level, as at the individual level, value indices have substantial predictive validity.

5. CONCLUSION

This article has emphasized the importance of studying values both at the level of individuals and at the aggregated level of nations. Evidence suggests substantial predictive validity at both levels. We can best explain cultural differences if we focus on both levels, while taking care to treat them separately. Difficulty in conceptualizing the interplay of personal and situational influences has bedeviled many areas of psychological investigation. The study of values does not provide a complete basis for understanding cultural differences, but it has proved one of the most promising initiatives. Progress has been made in identifying dimensions of variance, rather than studying single values, and in deriving these dimensions from theory.

As previously noted, nations do not necessarily provide the best basis for cross-cultural comparisons. Many people believe that there are major regional and/or ethnic differences in values within their country. Multiple samples from single countries are needed to test the validity of these beliefs. Even when within-nation differences are found, it is important to compare their magnitude with cross-national differences. Significant within-nation differences in SVS scores in China, Japan, and the United States are dwarfed by the much larger differences found between these nations. However, value diversity is likely to be greater in some nations than others. High SVS national value consensus is predicted by high socioeconomic development but low political democratization.

The world is currently experiencing rapid social change, driven by globalization, economic development, migration, and the information revolution. These changes may well entail marked shifts in the values endorsed by successive generations. Traditional ways in which values are transmitted from one generation to the next may be under threat. Consistency between results obtained by Hofstede and Schwartz suggests that these changes may be relatively gradual, but researchers have only recently started to map value changes systematically.

See Also the Following Articles

Attitudes ■ Ideological Orientation and Values ■ Interpersonal Conflict

Further Reading

Chinese Culture Connection. (1987). Chinese values and the search for culture-free dimensions of culture. *Journal of Cross-Cultural Psychology*, *18*, 143–164.

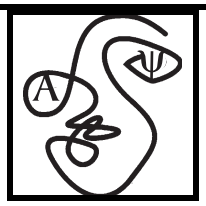
Hofstede, G. (2001). *Culture's consequences: Comparing values, behaviors, institutions and organizations across nations*. Thousand Oaks, CA: Sage.

Inglehart, R. (1997). *Modernization and post-modernization: Cultural, political and economic change in 43 societies*. Princeton, NJ: Princeton University Press.

Schwartz, S. H. (1992). Universals in the structure and content of values: Theoretical advances and empirical tests in 20 countries. *Advances in Experimental Social Psychology*, *25*, 1–65.

Smith, P. B., & Schwartz, S. H. (1997). Values. In J. W. Berry, M. H. Segall, & C. Kagitcibasi (Eds.), *Handbook of cross-cultural psychology* (2nd ed., Vol. 3, pp. 77–118). Needham Heights, MA: Allyn & Bacon.

Triandis, H. C. (1995). *Individualism and collectivism*. Boulder, CO: Westview.



Vocational Assessment in Schools

Edward M. Levinson, Angela B. Christy, and Marcy R. Maus

Indiana University of Pennsylvania, Indiana, Pennsylvania, USA

1. Introduction
2. Best Practices
3. Roles for Professionals
4. Conclusion
   Further Reading

GLOSSARY

transdisciplinary school-based vocational assessment (TVA) A comprehensive assessment conducted within a school setting whose purpose is to facilitate educational and vocational planning in order to allow a student to make a successful adjustment to work, post-secondary education, and community living.

transition Movement of a student from high school to post-secondary training and from home to more independent living.

School-based vocational assessment should incorporate a transdisciplinary approach that includes school and community professionals in the evaluation and assessment of the vocational and educational needs of students, especially those with disabilities. Several steps are involved in program development, including the following: planning and development, implementation, and evaluation and improvement. Once developed, programs should ensure that assessments are conducted at three levels: elementary, middle school/junior high, and high school. As the students progress through school, different skills are assessed with different assessment techniques, and students are provided with the skills and training necessary to make a successful transition from school to work and to community living.

1. INTRODUCTION

1.1. Importance

Students continue to progress through schools expecting that some job or career awaits them upon graduation from high school. They attend classes to enhance their skills in defined subject areas but ultimately fall short when it comes to possessing the skills necessary to obtain and uphold a gratifying job.

According to the U.S. General Accounting Office, one-third of young workers ages 16–24 do not have the skills they need to perform entry-level, semiskilled jobs. Additionally, nationwide studies conducted by the National Career Development Association and the Gallup Organization suggest that adults who consciously plan and prepare themselves for a job comprise only one-third of those currently working. Consequently, the remaining two-thirds of employed adults enter their jobs as a result of chance, the influence of others, or because they took the only job available to them. Most important, only half of these workers are content with their jobs.

In particular, individuals with disabilities are at especially high risk for failure or dissatisfaction with

©2004 Elsevier Inc.
All rights reserved.

their jobs. When compared with their nondisabled peers, individuals with mild disabilities experience a higher rate of unemployment and underemployment, lower pay, and greater dissatisfaction with their jobs. Hence, this population of students is especially in need of vocational assessment services.

Moreover, there is a high drop-out rate among students with disabilities; as a result, they have not received the vocational training necessary to make a smooth transition from school to work. Studies comparing special education drop-out rates with control group drop-out rates (normative data) have consistently demonstrated that students with disabilities leave school more often than students without disabilities. Students with learning disabilities and emotional disabilities are at an especially high risk of dropping out of school. Some studies have reported drop-out rates exceeding 40 and 50%, respectively, for these populations.

This high drop-out rate may be attributed to several factors. Many schools focus their curriculum on college-bound students rather than those who plan to enter the workforce following high school. Statistics indicate that only 15% of freshmen entering high school graduate and obtain a degree from a 4-year college within 6 years of their high school graduation. As a result, focusing resources on college-bound students leaves the majority of non-college-bound students lacking in the skills necessary for success in the workforce. The "21st Annual Report to Congress" reported that students with disabilities were more likely to stay in school and even more likely to be employed after high school if they were provided with vocational training.

The Carl D. Perkins Vocational Education Act and the Individuals with Disabilities Education Act (IDEA) both address vocational assessment in schools. The Perkins Act stipulates that parents must be informed about vocational education available for their children 1 year before they are eligible. The act also requires that parents be provided with information about vocational education, such as eligibility requirements and details about the programs offered. Once enrolled, students should be assessed in several areas, including interests, abilities, and special needs as well as any other special services needed to assist in the transition from school to work. IDEA mandates that services aiding in the transition process be established for students with disabilities to ensure a successful transition from school into work and the community. This law requires that plans for transitioning be initiated by age 14.

The process of establishing vocational assessment programs requires effective planning and a structured, team-oriented approach. The focus of this article is on the role of school psychologists in vocational assessment since they are the applied psychologists working in schools. School psychologists play an important role in providing transition services to students. In order to effectively and competently conduct a comprehensive vocational assessment, school psychologists need to be aware of career development theory and be knowledgeable about interest and aptitude assessment as well as other empirically based vocational assessment practices.

1.2. Vocational/Career Development Theory

Vocational/career developmental theory should be the basis for developing school-based vocational assessment programs. Vocational development theory also assists educators in making developmentally appropriate recommendations for students.

Vocational developmental theory suggests that individuals progress through stages. The growth and exploration stages span the school years and are each characterized by three substages. Within the growth stage, children progress through the fantasy substage (ages 0–10 years), in which they use their imagination to take on different career roles; the interest substage (ages 11 and 12 years), in which they consider different careers and areas of interest; and the capacity substage (ages 13 and 14 years), in which they become aware of career demands and the academic training required of careers following high school. The exploration stage consists of the tentative, transition, and trial substages. The tentative substage (ages 15–17 years) is a period during which the adolescent considers various career options and chooses a tentative career goal. During the transition substage (ages 18–21 years), an individual chooses a career and works toward acquiring the skills needed to enter into that career. The trial substage (ages 22–24 years) involves the attainment of a job that fits their career choice. One purpose of vocational assessment is to make certain that students are successfully progressing through these vocational developmental stages and to determine whether students require assistance in reaching the goals associated with each stage.

1.3. Skills Needed for Employment

In addition to understanding various aspects of one's job, there are several other characteristics that promote success in the work environment. For instance, employees

need to work well with others (social skills) and to make informative/effective decisions (decision-making skills) in the workplace.

An inclusive vocational assessment evaluates an individual's level of functioning in several domains. It should include measures of one's psychological, social, educational, physical, and vocational abilities. Reviewing records, conducting observations, and administering paper-and-pencil tests assist psychologists in gathering relevant information in these domains. In addition, performance tests, work samples, and situational assessments are also used in a comprehensive vocational assessment program. A performance test focuses less on language and more on an individual's ability to perform a specific task related to a job. Work samples consist of three phases (demonstration, training, and assessment) and expose the student to natural job responsibilities. Situational assessment measures an individual's interests, abilities, and work habits in actual or contrived work environments.

2. BEST PRACTICES

2.1. Systemic vs Individual Planning and Program Development

A school-based vocational assessment program is designed to meet the educational needs of all students, especially students with disabilities. Development and implementation of a program should be composed of three phases: planning and development, implementation, and evaluation and improvement.

Planning and development is composed of several steps. First, a representative team of school and community personnel likely to be involved in the assessment process or otherwise engaged in the program is formed. Second, a needs assessment is conducted, which involves identifying school and community resources/services to include in the program as well as recognition of possible complications associated with vocational planning. Following a needs assessment, an assessment model should be developed that incorporates local resources and reduces possible obstacles to programming. Fourth, formal interagency agreements and action plans should be arranged between school and community agencies. The purpose of these agreements is to identify what services will be provided by each agency, which students will receive services, and when services will be initiated. Last, if it is necessary to secure additional funds, the team finds sources of financial support for the program.

The implementation phase should begin with the hiring of a vocational assessment program coordinator. The coordinator is responsible for supervising the implementation of the program and locates sites where evaluations will be conducted. In addition, this individual should develop a procedures manual, decide on and purchase necessary materials, and train involved personnel. He or she is also responsible for conducting in-service workshops to familiarize school and community personnel with the assessment program. The final stages of the implementation phase involve a pilot testing of the program with a target group of students followed by revisions and full implementation of the program.

The first two steps of the evaluation phase acknowledge areas of the program in need of evaluation and identify standards for the evaluation process. Then, a program evaluator is hired to conduct the evaluation. Lastly, the vocational assessment team plans and implements necessary improvements. Program evaluations and revisions are conducted periodically to improve program functioning.

2.2. Assessment Models

There are several levels of assessment in school-based vocational assessment programs. Each level has a different purpose, answers different questions, uses different evaluation techniques, and gathers different kinds of information. The type and reason for assessment at a given level are based on developmental objectives and referral questions. Programs are generally composed of two or three levels of assessment. Level 1 assessments are conducted during the elementary school years. These evaluations focus on the child's understanding of self, interpersonal skills, and decision-making skills. In addition, the assessments include vocational and career exploration activities and are intended to improve self-awareness. Level 2 assessments occur in the middle school or junior high school years. The evaluation focuses on student interests, abilities, work habits, and career maturity. The vocational assessment team consults student records, conducts interviews and observations, and evaluates students with standardized norm-referenced assessment tools. The goal of level 2 assessments is to maintain and encourage career exploration and to assist students in formulating tentative career goals. Based on the results, vocational training programs are recommended for students based on their interests and abilities. In addition, program adaptations and

support services are recommended to increase student success in their selected program. Level 3 assessments are conducted in high schools. In order to evaluate high school students, the team utilizes a variety of assessment tools including work samples and situational assessments to determine the training needed to attain post-high school education or employment goals. Evaluation at this level also assists in determining the skills an individual needs to make a successful transition from school to work, post-secondary educational training, and community living.

Transdisciplinary school-based vocational assessment (TVA) is one program model that can be used as a framework for the planning and establishment of a school-based program. The model is equally applicable for students with or without disabilities. TVA uses assessments and interventions as well as involvement from school and community professionals to plan and develop a program. The model consists of four phases and involves two levels of assessment. Phase 1 is the development of a school-based vocational assessment program, in which the school and community team plans, organizes, and implements the program. During phase 2, students in grades 6–9 are administered a level 1 vocational assessment to determine educational and vocational goals, which are then incorporated into students' Individual Education Plans (IEPs). The data from the evaluation are also used to assist team members in identifying vocational training opportunities and future residential living options, making necessary curriculum changes, and connecting students to available school or community resources. Phase 3, designed for students in grades 10–12, incorporates a level 2 vocational assessment that is designed to assist the team in making revisions to a student's program and IEP. Phase 4 is designed to monitor training and placement in a job, a post-secondary school, and/or a residential living facility. Follow-up and continual support from the team are provided during this phase in order to aid in a successful transition from school to work and community living.

3. ROLES FOR PROFESSIONALS

There are several roles that the school psychologist can assume when developing or implementing a vocational assessment program. These vary based on existing vocational services in the district and the expertise of other school personnel involved in the assessment process. Ideally, the school psychologist is one of several school- and community-based professionals on the vocational assessment team. As a team member, he or she collects information about a student's ability, academic functioning, and social performance and considers its implications from a vocational perspective. Additional measures of vocational interests and aptitudes, an evaluation of student work, an observation in the work environment, and student interviews may also be conducted when assessing the vocational needs of an adolescent. The school psychologist may also use his or her knowledge of measurement principles to aid in the selection of appropriate assessment tools and techniques to be used in the evaluation process. Lastly, the school psychologist's understanding of learning theory and adolescent psychology allows him or her to serve as a consultant for others involved in the vocational assessment and planning process.

If a school-based vocational assessment program has not already been developed, the school psychologist may take on the responsibility of developing or assisting in the development of a program. Since a comprehensive vocational assessment program involving a team of professionals may not be possible in some schools, the school psychologist may be the sole individual involved in conducting vocational assessments.

4. CONCLUSION

Unfortunately, conventional vocational and transition assistance has not resulted in desirable outcomes for students with disabilities. The high unemployment and underemployment rates as well as the high drop-out rates associated with this population suggest a need for a change in focus from an academic to a vocational orientation. Vocational assessment programs in schools are valuable in preparing students with the skills necessary to acquire and maintain a satisfying job and to make a successful transition to adult life. Vocational assessment supplies team members with the information necessary to help students develop educational and vocational goals and assists students in receiving the training necessary for a successful transition from school to work and community living. The program allows for active participation from school and community professionals, and it includes an evaluation of the student's psychological, academic, physical, vocational, and social functioning.

School psychologists play a vital role in the vocational assessment process. Their psychometric expertise is useful in the selection of appropriate evaluation tools, and their knowledge of measurement ensures the accurate interpretation of data. Furthermore, their

understanding of learning and behavior theory and adolescent psychology is valuable in the development and implementation of the program.

See Also the Following Articles

Career Counseling ■ Holland's Theory (Vocational Personality Types) ■ Indecision, Vocational ■ Vocational Interests

Further Reading

Brown, D., & Minor, C. (Eds.). (1989). *Working in America: A status report*. Alexandria, VA: National Career Development Association.

Clark, G. B. (1998). *Assessment for transitions planning: A guide for special education teachers and related service personnel*. Austin, TX: PRO-ED.

Hoyt, K. B., & Lester, J. L. (1995). *Learning to work: The NCDA Gallup survey*. Alexandria, VA: National Career Development Association.

Levinson, E. M. (1993). *Transdisciplinary vocational assessment: Issues in school based programs*. Brandon, VT: Clinical Psychology Publishing.

Levinson, E. M. (1995). Best practices in transition services. In A. Thomas, & J. Grimes (Eds.), *Best practices in school psychology III*. Washington, DC: National Association of School Psychologists.

Levinson, E. M. (1998). *Transition: Facilitating the post-school adjustment of students with disabilities*. Boulder, CO: Westview.

Levinson, E. M., & Brandt, J. (1997). Children and career development. In G. Bear, K. Menke, & A. Thomas (Eds.), *Children's needs II*. Washington, DC: National Association of School Psychologists.

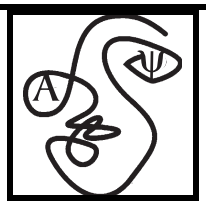
Levinson, E. M., & Murphy, J. (1999). Transition and school psychology. In S. DeFur, & J. R. Patton (Eds.), *Transition and school-based services: Interdisciplinary perspectives for enhancing the transition process*. Austin, TX: PRO-ED.

Phelps, L. A., & Hanley-Maxwell, C. (1997). School-to-work transitions for youth with disabilities: A review of outcomes and practices. *Review of Educational Research*, *67*, 197–226.

Power, P. W. (2000). *A guide to vocational assessment* (3rd ed.). Austin, TX: PRO-ED.

Smith, F., Lombard, R., Neubert, D., Leconte, P., Rothernbacher, C., & Sitlington, P. (1995). *Position paper of the Interdisciplinary Council on Vocational Evaluation and Assessment*. Available at: http://www.vewaa.org/council.html.

Super, D. E. (1957). *The psychology of careers*. New York: Harper.



Vocational Interests

Fred H. Borgen

Iowa State University, Ames, Iowa, United States

1. History: Eighty Years of Interest Measurement
2. Two Types of Scales
3. Major Interest Inventories
4. Quality and Use of Interest Measures
5. Trends in Interest Measurement
6. Conclusions

Further Reading

GLOSSARY

career self-efficacy One's belief that one can successfully perform the behaviors required in a particular career-related domain, such as mathematics or sales.

contrasted-groups scale A scale that is built by identifying items responded to differently by a criterion group and by a contrast group (or norm group). The best known examples are the clinical scales of the MMPI and the occupational scales of the Strong Interest Inventory.

Holland's hexagon John Holland suggested that there are six major interest dimensions or types that can be arrayed in order around a hexagon, with interests that are nearer on the hexagon (e.g., investigative and artistic) being more similar psychologically and interests that are across the hexagon (e.g., investigative and enterprising) being the most different psychologically.

homogeneous content scale A scale designed to have items of similar content that correlate with each other. All interest inventories measuring Holland's RIASEC dimensions or types use content scales that represent the broad content within each type. There are also more narrowly focused content scales, such as the Basic Interest Scales of the Strong Interest Inventory. Content scales have high internal consistency reliability.

occupational scale First created by Strong, a contrasted-groups scale containing items empirically differentiating people in an occupation from a general group of people; high scores show relative similarity to the distinctive likes and dislikes of people in that occupation. Such scales typically have a heterogeneous mix of items; hence, their linguistic content is difficult to interpret.

RIASEC Acronym for Holland's six interest dimensions or types: realistic, investigative, artistic, social, enterprising, and conventional.

vocational interest An emotional reaction to a career-related stimulus, ranging from distinctly liking to distinctly disliking.

Beginning more than 80 years ago, the assessment of vocational interests is one of the great success stories of applied psychology. Millions of vocational interest inventories are used every year, usually by young people exploring career and educational options and choices. Increasingly, they are also used by more experienced workers who are evaluating their current situations or exploring career transitions. Interest inventories have a long and vigorous history in psychology, marked by intensive and groundbreaking research, technical innovation, theoretical consolidation, and adaptation to changing societal needs. Today, interests play as important a role as ever in assessing individuality, especially in concert with other domains, such as abilities, personality, and self-efficacy. Hundreds

©2004 Elsevier Inc.
All rights reserved.

of vocational interest inventories exist, but the following five (in alphabetical order) can be considered the most widely used and well researched: the Campbell Interest and Skill Survey, the Kuder Occupational Interest Survey, the Self-Directed Search, the Strong Interest Inventory, and the UNIACT interest inventory.

1. HISTORY: EIGHTY YEARS OF INTEREST MEASUREMENT

Two dominant themes in the long history of vocational interests are E. K. Strong's resolute empiricism and John Holland's seminal theory. Strong mapped the interests of people in various occupations, creating a large and still growing database. Holland focused the interpretation of increasingly complexity of interest data with a theory that makes interests more understandable and accessible. The traditions originated by Strong and Holland continue to profoundly influence research and reflection about the meaning and measurement of vocational interests.

1.1. Strong's Empirical Success

The first widely used interest inventory appeared in 1927 with Strong's construction of the Vocational Interest Blank at Stanford University. Strong's innovative approach was intensely empirical, constructing occupational scales consisting of those interest items that men in a specific occupation answered differently than "men-in-general." Strong was passionate in cataloguing interests, often following the careers of people over several decades. His classic 1943 book best exemplifies his data-driven approach to interest assessment. Strong's insights in this book presage developments that might be considered cutting edge today. Despite the limits of the technology of his day, Strong used pioneering methods such as factor analysis to examine the structure of interests. From this he constructed a classification of occupations that foreshadowed Holland's famous RIASEC hexagon.

1.2. Holland's Influential Theory

Holland has been the leading theorist in vocational psychology during the past 40 years. His novel approach asserts that interests are an expression of personality. He seems to say that interests and personality are synonymous, yet at the same time he is well-known for a widely used interest inventory, the Self-Directed Search. Moreover, his interest inventories are similar in many ways to others that do not claim to measure personality. A current appraisal might be that interests and personality are treated as separate domains by most theorists and counselors, although there is evidence that they overlap in many ways.

Holland proposed that there are six primary dimensions of vocational interests, characterized by his famous RIASEC acronym: (i) Realistic interests involve building and mechanical activities; (ii) investigative interests encompass science and mathematics; (iii) artistic interests include art, music, dramatics, and writing; (iv) social interests entail helping and teaching; (v) enterprising interests involve managing, selling, and persuading; and (vi) conventional interests include organizing and evaluating data. Holland derived these dimensions from his own empirical work and from the research literature, including Strong's 1943 book. These dimensions quite adequately map the primary broad dimensions of vocational interests.

Holland and his associates' insight of the early 1970s that the RIASEC dimensions (or types) array themselves on a hexagon made Holland's approach very popular in both theory and applications. The hexagonal arrangement permits a Holland calculus based on relations around the hexagon. Distance on the hexagon translates into psychological distance: For example, scientists (I types) are across the hexagon from salespeople (E types), and these people differ on an array of psychological dimensions, including interests. The RIASEC hexagon provides an efficient summary of the major dimensions of interests and their relations to each other. Therefore, a number of popular interest inventories have incorporated the RIASEC dimensions. Many of these inventories measure level of interest on these six dimensions, and many use the Holland themes as an organizing principle. Holland also generates a number of theoretical propositions from the hexagonal relationships, and these have stimulated considerable research during the past four decades.

1.3. Interest as a Construct often Poorly Defined

In one sense, interest measurement has been a victim of its own success. Strong's empirical occupational scales garnered much respect and utility so their conceptual limitations were forgiven. Conceptual definition of interests did not frame Strong's occupational scales.

The scales were more in the tradition of dustbowl empiricism, with an occupational scale defined by the items that empirically differentiated the people in that occupation. Consequently, we do not have a well-developed theory of vocational interests that explains their meaning as constructs and their origins and action. Rather, there are nascent developments of such a theory. Recent work, exemplified by that of Savickas and Spokane, has attempted to imbue interest measures with enhanced meaning. Holland's theory is a major step toward understanding interests conceptually. The rise of content scales is another major step toward scales with direct meaning and interpretability. Finally, social cognitive career theory is a third major step toward understanding interests in the context of other variables, such as career self-efficacy, and developmental processes.

2. TWO TYPES OF SCALES

2.1. Occupational Scales (Contrasted Groups)

Constructing occupational scales by using the contrasted-groups method is Strong's most respected innovation and most enduring legacy. Occupational scales have remained the centerpiece of the Strong Interest Inventory for nearly eight decades. They are the basis for the lasting respect given the Strong inventory for its intense and careful empiricism tracking the likes and dislikes of satisfied workers in national samples of specific occupations. The Strong inventory has been routinely revised by collecting large samples of men and women in diverse occupations.

In the latest revision of the Strong inventory, Harmon *et al*. followed the steps Strong outlined for constructing occupational scales. As had been done from the beginning, separate occupational scales were created for women and men based on the somewhat different interests of women and men in general and sometimes within specific occupations. Thus, Harmon *et al*. created a new occupational scale for female physicists and a new scale for male physicists. Using the female physicists as an example, the steps can be described. The rationale for an occupational scale is to differentiate people in that occupation from a general reference group. High scores on a scale indicate that the person shares the likes and dislikes that differentiate people employed in that occupation from people in general. Thus, in this contrasted-group method, the responses of 348 female physicists were compared to those of other women on each of the 317 items in the 1994 Strong. Those items for which they differed—either in the like or the dislike direction by 27% or more—became the 56 items in the female physicist occupational scale. Some things that female physicists like more than other women are obvious, such as science and math. Women physicists also like some things that might not be intuitively obvious, such as mechanical activities and statistics. They also distinctively dislike some things more than most other women, namely business, fashion, and extroverted activities.

Because of its method of construction, and also because of the type of the people in most occupations, an occupational scale has some distinctive features that differ from many psychological scales. First and foremost, it has heterogeneous content—a mix of relative likes and dislikes that distinguish people in that occupation. Second, as a corollary to the heterogeneous content, an occupational scale is difficult to interpret in content terms. It can be interpreted in an important actuarial sense by saying that a high score signifies the person shares the distinctive likes and dislikes of people in that occupation. That is a very important piece of information but not one that conceptually disentangles the types of likes and dislikes that lead to a high or low occupational score. This problem led to the development of homogeneous content scales.

2.2. Homogeneous Content Scales

The first popular interest inventory to use homogeneous content scales was the Kuder Preference Record. This pioneering work by Frederic Kuder dates back to the 1930s. Today, content scales are at least a part of nearly every major interest inventory. Some, such as Holland's Self-Directed Search, use exclusively content scales and do not use occupational scales at all. Content scales are designed to have directly interpretable content in a specific domain, such as mechanical interests. Content scales are made up of items that tend to correlate in a general group of people, and thus, by definition, they are homogeneous. Content scales are similar to the factors that might be derived from factoring a set of interest items, but most inventories do not use this approach to create factor dimensions. Instead, content scales are often constructed to fit a rational or conceptual map of important and relatively independent interest domains. Often, items are included in an inventory to tap a particular domain, and an inventory could be revised

by adding items likely to create a new content scale. Content scales are typically built by examining item intercorrelation matrices of the tentative items to refine the item set to enhance homogeneity or internal consistency reliability of the scale. Holland's RIASEC dimensions are the best example of broad content scales, and the majority of popular inventories have measures of these six dimensions. Many inventories also have basic interest scales, which are content scales with more specific focus than the Holland dimensions. Thus, an inventory might have an artistic Holland scale yet also include basic interest measures of its components of fine arts, music/dramatics, and writing.

3. MAJOR INTEREST INVENTORIES

Five of the most popular interest inventories are the Campbell Interest and Skill Survey (CISS), the Kuder Occupational Interest Survey (KOIS), the Self-Directed Search (SDS), the Strong Interest Inventory (SII), and the UNIACT Interest Inventory. The SDS and UNIACT inventories represent the most simplified approach to interest measurement because their cores are primarily the six Holland RIASEC dimensions. The most complex inventories, because of the number and kinds of scales, are the Strong and the Campbell inventories. The Kuder inventory is somewhat less complex and represents an intermediate level of complexity. The utility of these inventories has been extensively supported by research. Interest scales effectively differentiate and predict important career behaviors, such as career choice and persistence, college major, and educational aspirations.

3.1. Campbell Interest and Skill Survey

Campbell's CISS is the newest of the major interest inventories. Earlier, Campbell was an important innovator with the Strong, so it is not surprising that the CISS is like the Strong in core ways: occupational scales, Holland-like scales, basic interest scales, and an organization by the Holland-like scales. It differs from Holland's scales in that it renames the RIASEC dimensions in more ordinary language and it adds an additional general content scale for adventuring. Even more distinctive in the CISS is the addition of self-ratings of skills to match the interest dimensions. For example, there is a scale for mathematics interests but also a parallel scale for mathematics skill. Such parallel scales are plotted together so that discrepancies between interests and skills in the same domain are highlighted. Campbell's use of separate skill ratings has been emulated in other assessment inventories.

3.2. Kuder Occupational Interest Survey

Kuder was one of the early pioneers of interest measurement. The KOIS is his approach to occupational interest scales, one that is quite different from Strong's. Kuder measured an individual's direct similarity to the average interests of people in an occupation, whereas Strong's occupational scales measure items that differentiate people in an occupation from those in a general group of other occupations.

3.3. Self-Directed Search

The SDS follows the Holland mantra for simplicity and ease of use. Its items are self-scored to yield scores on the six RIASEC dimensions. These scores are then translated into a RIASEC code that indicated the three highest scores in the profile.

3.4. Strong Interest Inventory

The SII, the oldest of today's inventories, has been revised nearly every decade to update the occupational scales that are the inventory's centerpiece. Since the beginning the SII has had a data-driven base, and it is widely viewed as the "most scientifically grounded" tool for interest assessment. The 1994 version of the SII contains 211 occupational scales constructed separately for women and men. Starting in the 1960s, the authors of the Strong added content scales—first the Basic Interest Scales, then the Holland General Occupational Theme Scales, and, recently, the Personal Style Scales. Beginning in 1996, the Strong has been available with the Skills Confidence Inventory, which measures confidence in each of the six RIASEC dimensions.

3.5. UNIACT Interest Inventory

Unlike the other four inventories, the UNIACT is not commercially available to the public. It is part of the assessment tools available to U.S. high schools through the American College Testing Program. It uses six content scales to measure the six RIASEC dimensions

and then maps the scores on a world-of-work map, which is an extension of Holland's hexagon.

4. QUALITY AND USE OF INTEREST MEASURES

Vocational interests show remarkably robust psychometric characteristics, presumably because the distinctive likes and dislikes of people are deep and often enduring. By young adulthood, most people are quite clear about their passions and aversions. Moreover, people in different educational and occupational settings are often starkly different in what they like and dislike. The following are examples from the Strong inventory norm group: 74% of 200 women airline attendants liked "looking at things in a clothing store" versus 33% of 200 women physicists. For these same women, 65.5% of the flight attendants said they disliked "calculus" as a school subject versus none of the physicists.

4.1. Reliability

Content scales are expected to have high item intercorrelations; this internal consistency reliability is usually expressed with Cronbach's alpha. Because interests are such robust phenomena, it is rather easy to develop interest scales with high reliability. Thus, the six Holland RIASEC scales in the 1994 Strong all have alpha reliabilities of 0.90 or more. Occupational scales are expected to be heterogeneous, so a different approach to reliability is appropriate. The typical approach is to examine test–retest reliability of occupational scales. Interests are usually quite stable over long periods of time. For example, people who liked mathematics as adolescents are likely to have the same feelings 30 years later.

4.2. Validity

There is a large literature on the validity of vocational interests for differentiating occupational and educational groups and for predicting career outcomes. Most of the references in the Further Reading section speak to such issues. Holland's observation that aspirants for occupations often resemble those who are in occupations leads to the probability that evidence for concurrent validity will map into evidence for predictive validity. This means that if we can identify the distinctive interests of adults currently in an occupation, such a cluster of interests will also longitudinally predict entry to such an occupation, or a similar occupation, for young people, or adult career changers, developing a career trajectory. This does not negate the special value of predictive studies, but it does convey the considerable value of concurrent validity studies.

4.3. Counseling Utility

Especially in the hands of a skilled career counselor, interest inventories can be useful in several ways. The best use will help clients better understand themselves and their vocational options and stimulate directions for action. Best practices in vocational assessment blend subjective information from counseling with objective information such as interest inventories to reveal the client's career story.

Broad scales, such as Holland's RIASEC scales, and specific scales, such as basic content scales, have different roles to play in counseling and in self-understanding. The broad scales are readily understood and readily organize broad career options. They are the best place to start the process of formulating a career story and charting directions for exploration. The specific scales then supplement this broad information when clients start refining their career story and their more precise paths for exploration. A broad enterprising scale points to a broad venue of influencing, yet more precise scales, such as merchandising and public speaking, speak distinctively to the options of retail sales versus politics. The specific scales are especially useful in identifying client strengths because there are a larger number of them and therefore they are more likely to show some high scores indicating areas of interest.

5. TRENDS IN INTEREST MEASUREMENT

The level of research about interests has been intense and lively in the past decade. Much work has been done to examine whether theories and measures of interests can be generalized across diverse groups. There is also a ferment of activity about how domains of interests relate to other domains of individuality such as personality and self-efficacy.

Recent research and theorizing have examined how interests are related to other domains of individuality, leading to speculation that there is a unified set of underlying traits linking interests, personality, and self-efficacy. This makes for exciting prospects for

both theory and applied assessment. For example, work in social cognitive career theory posits a developmental relationship whereby emerging self-efficacy leads to interest development. This causal proposition is stimulating much longitudinal research. For example, Nauta *et al*. found that the causal relationship was reciprocal, with interests leading to self-efficacy and conversely. Many interest inventories have a parallel set of measures of self-efficacy and there is burgeoning interest in their joint counseling use and theoretical implications.

Although Holland's position is that interests are expressions of personality, other writers have questioned the degree of overlap between the two domains. This stimulated a number of studies in the 1980s and 1990s to determine how interests and personality correlate with each other. A useful summary of this literature is provided by Larson *et al*.'s meta-analysis of 24 samples completing an interest inventory measuring Holland's Big Six dimensions and a personality inventory measuring the Big Five personality dimensions. From their results, it is evident that certain personality and interest dimensions are clearly related. Correlated 0.30 or more are personality openness and artistic interests, extroversion and enterprising interests, and extroversion and social interests. Also related are openness and investigative interests, and personality agreeableness and social interests. Such results demonstrate that there are clear links between certain broad dimensions of interests and broad dimensions of personality. The personality–interest links are even stronger and more meaningful when more specific dimensions are examined.

6. CONCLUSIONS

For more than 80 years, vocational interests have been studied and assessed with considerable success. Vocational interests are robust phenomena. The interests of people and occupations differ greatly and in theoretically meaningful ways. Several widely used interest inventories include high-quality measures with utility in counseling for building the client's career narrative. Vocational interests are the cornerstone of effective career counseling that generates options for exploration. Increasingly, it is apparent that other domains of individuality, such as personality and self-efficacy, are vitally linked to interests. A unified and integrated approach to all these domains will lead to enlightened theory and career counseling.

See Also the Following Articles

Career Counseling ■ Holland's Theory (Vocational Personality Types) ■ Indecision, Vocational ■ Vocational Assessment in Schools ■ Vocational Psychology, Overview

Further Reading

Betz, N. E., Borgen, F. H., & Harmon, L. W. (1996). *Manual for the Skills Confidence Inventory*. Palo Alto, CA: Consulting Psychologists Press.

Borgen, F. H., & Lindley, L. D. (2003). Optimal functioning in interests, self-efficacy, and personality. In W. B. Walsh (Ed.), *Counseling psychology and optimal human functioning* (pp. 55–91). Hillsdale, NJ: Erlbaum.

Campbell, D. (2002). The history and development of the Campbell Interest and Skill Survey. *Journal of Career Assessment*, *10*, 150–168.

Day, S. X., & Rounds, J. (1998). Universality of vocational interest structure among racial and ethnic minorities. *American Psychologist*, *53*, 728–736.

Diamond, E. E., & Zytowski, D. G. (2000). Vocational assessment. In C. E. Watkins, & V. L. Campbell (Eds.), *Testing and assessment in counseling practice* (2nd ed.). Mahwah, NJ: Erlbaum.

Donnay, D. A. C., & Borgen, F. H. (1996). Validity, structure, and content of the 1994 Strong Interest Inventory. *Journal of Counseling Psychology*, *43*, 275–291.

Donnay, D. A. C., & Borgen, F. H. (1999). The incremental validity of vocational self-efficacy: An examination of interest, self-efficacy, and occupation. *Journal of Counseling Psychology*, *46*, 432–447.

Fouad, N. A., & Zao, K. E. (2000). Meeting the revolution: Future trends in vocational assessment. *Journal of Career Assessment*, *8*, 403–409.

Harmon, L. W., Hansen, J. C., Borgen, F. H., & Hammer, A. L. (1994). *Applications and technical guide for the Strong Interest Inventory*. Palo Alto, CA: Consulting Psychologists Press.

Holland, J. L. (1997). *Making vocational choices* (3rd ed.). Odessa, FL: Psychological Assessment Resources.

Larson, L. M., Rottinghaus, P. J., & Borgen, F. H. (2002). Meta-analyses of Big Six interests and Big Five personality variables. *Journal of Vocational Behavior*, *61*, 217–239.

Nauta, M. M., Kahn, J. H, Angell, J. W, & Cantarelli, E. A. (2002). Identifying the antecedent in the relation between career interests and self-efficacy: Is it one, the other, or both? *Journal of Counseling Psychology*, *49*, 290–301.

Prince, J. P., & Heiser, L. (2000). *Essentials of career assessment*. New York: Wiley.

Savickas, M. L. (1999). The psychology of interests. In M. Savickas, & A. Spokane (Eds.), *Vocational interests: Their meaning, measurement, and counseling use* (pp. 19–56). Palo Alto, CA: Davies-Black.

Savickas, M. L. (2002). Career construction: A developmental theory of vocational behavior. In D. Brown (Ed.), *Career choice and development* (4th ed., pp. 148–205). San Francisco: Jossey-Bass.

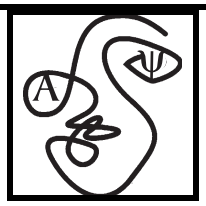
Savickas, M. L., & Spokane, A. R. (1999). *Vocational interests: Their meaning, measurement, and counseling use*. Palo Alto, CA: Davies-Black.

Savickas, M. L., Taber, B. J., & Spokane, A. R. (2002). Monograph: Convergent and discriminant validity of five interest inventories. *Journal of Vocational Behavior*, *61*, 139–184.

Walsh, W. B. (2002). Introduction: Exploring interests, self-efficacy, and goals. *Journal of Career Assessment*, *10*, 147–149.

Walsh, W. B., & Betz, N. E. (2001). *Tests and assessment* (4th ed.). Upper Saddle River, NJ: Prentice Hall.

Watkins, C. E., & Campbell, V. L. (Eds.). (2000). *Testing and assessment in counseling practice* (2nd ed.). Mahwah, NJ: Erlbaum.



Vocational Psychology, Overview

Mark L. Savickas

Northeastern Ohio Universities College of Medicine, Rootstown, Ohio, USA

1. Person-Environment Fit Model
2. Vocational Personalities and Work Environments
3. Career Model
4. Career Development
   Further Reading

GLOSSARY

career adaptability The readiness and resources for making career choices and negotiating career transitions.

planfulness An attitude that favors planning and preparing for anticipated future events.

psychological type A constellation of traits such as particular abilities and specific interests that characterize a group of individuals.

psychosocial A relationship or interaction between an individual and a social group.

skilling Engagement in training and education that prepares one to enter a particular occupation.

trait An underlying dimension that structures and explains recurring uniformity in clusters of behavior emitted by an individual.

When asked what characterizes mental health, Freud reputedly answered, "love and work." Although most psychologists prefer to focus on relationships, work also provides a fruitful focus for the study of personality and human development. This assertion is borne out in vocational psychology, the discipline in applied psychology that concentrates on how individuals experience their work, occupations, and careers. The basic unit of study in vocational psychology is "vocational behavior," that is, the responses an individual makes in choosing and adapting to an occupation. In addition to studying vocational responses to occupational stimuli, vocational psychologists also study the development of vocational behavior over the lifespan, from the growth of occupational aspirations during childhood to the disengagement from the world of work during old age. The overview of vocational psychology presented herein explains the structure and function of the discipline, particularly emphasizing its concern with person-environment fit and career development and introducing its major constructs.

1. PERSON-ENVIRONMENT FIT MODEL

The study of vocational behavior emerged as a scientific discipline at the beginning of the 20th century as industrialized countries across the globe saw increasing numbers of workers leave their farms to seek employment in the large factories that dominated modern cities. These companies assumed a bureaucratic form and a hierarchical organization. This organizational structure of ranked positions changed workers from self-employed individuals who performed a craft for their community to employees who performed just

©2004 Elsevier Inc.
All rights reserved.

some of the specialized tasks required to bring a product to market. Employers began to create jobs composed of a few circumscribed tasks. These factory jobs required only a small set of skills and interests, so employers sought to fill these jobs with employees who clearly displayed the required characteristics.

Upon arrival in the metropolis, with their immediate family in tow, individuals from the countryside and from different countries faced the task of finding employment. Unaccustomed to life in the city, with its factories and specialized occupations, many of these newcomers needed help in relating their talents to available occupations. Providing this vocational guidance became the mission of philanthropic organizations that helped newcomers to the city choose a job and adapt to urban life. Guidance workers in institutions such as the YMCA and settlement houses used a three-step model to match people to occupations. First stated by Parsons (1909), the person-environment (P-E) model consists of helping individuals increase self-knowledge and gather occupational information, and then rationally match the two to select a fitting occupation. Advocates of the model for P-E fit assert, with ample evidence, that a good match of a worker's abilities and interests to a job's requirements and rewards leads to important outcomes. The organization gets a satisfactory worker, one who is productive and reliable, and the successful worker earns economic rewards and feels job satisfaction.

The primary outcomes of fitting workers to work, that is, occupational success and job satisfaction, may result in an important secondary outcome—career stability. After leaving school, individuals usually progress through a sequence of trial positions in a process of zeroing in on a more-or-less permanent job. This sequence of trial jobs can entail horizontal movement to lateral positions in the hierarchy or vertical mobility up the organizational ladder. Ideally, each new job should improve correspondence between the individual's resources and the occupation's requirements, leading to more effective functioning. Once a good match has been made, individuals usually remain in that congruent position, enjoying a relatively stable work life. The length of this period of stable employment is the major variable in comprehending the wide range of work histories that individuals report. At one end of the continuum, some individuals find a fitting position, remain in it for 30 years, and then retire from it. At the other extreme, some individuals never stabilize for more than a year or two, having as many as 25 jobs in their career. A more typical work history finds workers remaining in a position for some significant period, say 5 to 7 years, before moving to and restabilizing in another position, repeating this process four to eight times during their work lives. Various career patterns are thus characterized by the sequence and duration of work positions occupied by an individual. Using variations in tenure, researchers have identified a set of common career patterns that accommodate the work histories of most individuals. The career patterns experienced by large groups of workers have been examined to identify how occupations relate to each other and to describe the vocational personality types that populate particular occupations. These observations became systematized as researchers studied individual differences in vocational personalities and similarities in work environments.

2. VOCATIONAL PERSONALITIES AND WORK ENVIRONMENTS

Development of the psychotechnology for implementing the person-environment paradigm was spurred by World War I, as the methods devised for matching soldiers to military positions were adapted to civilian occupations. Two lines of research dominated that era, and remain dominant today. The first line concentrates on success and the second on satisfaction. Efforts to predict success led both to the development of tests that measure individual differences in abilities, aptitudes, and skills and to the assessment of the ability patterns that characterize workers in various occupations. Efforts to predict satisfaction led both to the development of inventories that measure individual differences in vocational interests and work values and to the assessment of the interests reported by satisfied workers employed in various occupations. Vocational guidance counselors match individuals to fitting occupations by comparing the abilities and interests that characterize different occupations to the profiles individuals produce on ability tests and interest inventories. Of course, personnel officers use this same approach in employee selection. So, the P-E theory and its associated techniques can be used to benefit the employee, the employer, and the community.

While industrial-organizational (I/O) and vocational psychologists both appreciate the roles that abilities and interests play in shaping careers, they differ in the relative emphasis they place on these critical variables. Industrial psychology, with its focus on an organization's productivity and employee success,

emphasizes ability tests and occupational ability patterns. In contrast, because it concentrates on the perspective of the individual worker and job satisfaction, vocational psychology emphasizes research on work motivation. Although some attention has been, and continues to be, paid to motivational variables such as personal needs and work values, the majority of vocational psychology's research on motivation has concentrated on vocational interests.

2.1. Vocational Interests

Interest denotes a state of consciousness characterized by a readiness to respond to particular environmental stimuli, such as objects, activities, and people. When activated, interests prompt selective attention that narrows the perceptual field to concentrate on the stimulus. This attention is accompanied by a pleasant feeling and an evaluation of liking that steers goal-directed behavior and maintains persistent effort to fulfill some personal need or value. As a personality trait, vocational interests denote a homogeneous group of specific interests that form a consistent, persistent, and stable dispositional response tendency which increases one's readiness to attend to and act upon a particular group of environmental stimuli. This orientation shows habit (or absolute) strength in how much stimulation is required to activate it. Habit strength can be assessed by behavioral analyses and interest autobiographies that reveal the ease and frequency with which an interest is initiated, as well as the duration for which it extends and the length of time for which it persists. A disposition shows relative strength in activity preferences, that is, competition with other interests for behavioral expression. Relative strength of interests can be measured with interest inventories.

In addition to being viewed as a personality trait, vocational interests may also be conceptualized as a psychosocial construct, psychosocial meaning the interrelation of individual and society in human development, especially connoting the influence of social factors on an individual's thoughts and actions. From this perspective, vocational interests denote a complex, adaptive effort to use one's environment to meet needs for survival and self-realization. Vocational interests expedite P-E interactions by uniting subject, object, and behavior into a vital relationship, a relation manifested in actions that satisfy needs, fulfill values, foster self-development, enhance adaptation, and substantiate identity.

The dominant figure in the study of vocational interests during the 20th century was E. K. Strong, Jr., who in the 1920s devised a method for measuring an individual's vocational interests and a scoring system for determining empirically how well those interests resemble the interests of workers employed in a number of different occupations. Strong's inventory, with continuing updates, remains one of the more popular tools used in vocational psychology research and in career counseling practice. The current version of Strong's inventory incorporates two additional types of interest scales that augment his empirically keyed occupational scales (with heterogeneous items) for measuring degree of resemblance to different occupational groups. The first set of scales was added to Strong's inventory in the 1960s; they were basic interest scales composed of homogeneous items that measure expressed vocational interests. Scores on these rationally constructed scales are interpreted as a direct indication of an individual's vocational interests, in contrast to Strong's empirically constructed scales of heterogeneous items which indicate occupational interests by indexing resemblance to workers in different occupations. In the 1970s, a second type of scale was added to Strong's inventory, one that measured resemblance to the six prototypes in Holland's theory of vocational personality types.

2.2. Vocational Personality Types

Based on factor-analytic studies of interests and research with Strong's inventory, Holland logically deduced six types of vocational personalities and devised a heuristic theory to describe vocational behavior and organize work histories. The six prototypes in Holland's theory are each characterized by a distinct syndrome of interests, competencies, and activities.

1. *Realistic* (R) types report outdoor and mechanical interests, prefer to work with animals and machines, enjoy the role of doer, display physical competencies, like leisure pursuits involving physical skill and challenges, and admire role models such as athletes and adventurers. They might be heard often saying, "just do it."
2. *Investigative* (I) types report scientific interests, prefer to work with ideas, enjoy the role of thinker, display intellectual competencies, like leisure pursuits involving reading and researching, and admire role models such as scientists, inventors, and detectives. They can often be heard saying, "let's explore it."
3. *Artistic* (A) types report artistic, literary, and musical interests, prefer to work with feelings, enjoy the role of creator, display aesthetic competencies, like leisure pursuits involving self-expression and appreciation of

concerts, theaters, and museums, and admire role models such as artists, composers, writers, and performers. Often, they can be heard saying, "let's create it."

4. *Social* (S) types report social interests, prefer to work with people, enjoy the role of helper, display communication competencies, like leisure pursuits involving conversation and social gatherings, and admire role models such as teachers and social workers. They can often be heard saying, "let's talk about it."

5. *Enterprising* (E) types report sales and managerial interests, prefer to work with opinions, enjoy the role of leader, display persuasive competencies, like leisure pursuits involving travel and politics, and admire role models such as public officials, military officers, and corporation presidents. Often they could be heard saying, "make it happen" or "make it so."

6. *Conventional* (C) types report clerical and business interests, prefer to work with data and records, enjoy the role of member, display organizing competencies, like leisure pursuits involving collecting and organizing, and admire role models such as teams, organizers, and historians. They can often be heard saying, "keep it going."

Holland subsequently elaborated six types of environments, or ecological niches, that correspond to each of the six vocational personality types. Along with his colleagues, Holland then used his typology of environments to construct an occupational classification system as well as to classify other behavioral settings such as college majors, universities, leisure activities, and hobbies. Researchers type a behavioral setting by identifying the modal vocational personality types that populate it. Alternatively, one can type an environment by observing its contents and activities. For example, an automobile parts store with mechanical parts lining its neat shelves is a realistic and conventional environment, whereas an elementary school with children discussing their feelings is a social and artistic environment. Holland's hexagonal scheme for portraying the world of work has been compelling in vocational research and pragmatic in career counseling practice.

2.3. Consistency and Congruence

The six types of vocational personalities and work environments constitute the primary elements in Holland's theory. In due course, he elaborated four secondary constructs, starting with consistency. Holland determined the relative frequency of type patterns and used this information to order the six types around a hexagon, starting with R in at the upper left node, and followed clockwise by I, A, S, E, and C. Using the RIASEC ordering, researchers and practitioners can determine the degree of internal consistency for a vocational personality or work environment. Consistent types appear next to each other. For example, the most frequent secondary type for an individual with a primary type of artistic is investigative, which precedes A in the RIASEC ordering, or social, which follows I in the RIASEC ordering. Inconsistent types such as artistic and conventional are opposite each other; and, intermediate types such as artistic and enterprising are separated by one other type. The theoretical elegance and practical utility of Holland's research on the structure of interests has been augmented by other researchers who have elaborated the hexagon first by embedding it into an interest circle and then converting the circle into an interest globe.

In addition to the construct of consistency, which deals with the intrapersonal organization of interests, Holland also elaborated his theory by conceptualizing how people behave in environments. In using the same construct language to type both people and environments, Holland made it clear to even casual observers that adaptive fitness is optimal when people work in congruent environments. Simply stated, the construct of congruence proposes the idea that corresponding matches of person to environment produce success, satisfaction, stability, and mental health. This idea enabled Holland to show, for example, that an individual who is ISA, say a psychologist, would be congruent in an ISA work environment, say working as a professor, but incongruent in a REC environment, say working as a roofer.

To further characterize each type of person and occupation, and more importantly the match between them, Holland assesses degree of differentiation or intensity of interests. Marked differentiation indicates a specialist with narrow interests, whereas limited differentiation indicates a generalist with broad interests. A final construct, vocational identity, refers to how clear a picture an individual has of his or her own type and how many distinct behavior situations occur in a work environment. Other things being equal, Holland proposes that it is easier to predict work adjustment for an individual who works in an organization with few distinct behavioral situations and who displays a clear identity with markedly differentiated interests. While still the dominant paradigm in vocational psychology, the individual differences view of occupations was augmented in the 1950s by a second grand model, namely, the individual development view of careers.

3. CAREER MODEL

Career is the development of vocational behavior over time or, according to the *Oxford English Dictionary*, "a person's course or progress through life." The career model, which concentrates on the development of vocational behavior, shares an interest with the P-E model in being concerned with how individuals fit into occupations at a particular point in their lives, yet prefers to emphasize how individuals fit work into their lives from childhood to retirement. Instead of the occupational stimulus-vocational response (S-R) paradigm for studying vocational behavior, career studies use a response-response (R-R) paradigm. Vocational behavior, or response, remains the basic unit of study, but instead of studying differences in vocational behavior between individuals, the career perspective concentrates on changes in vocational behavior by the same individual across time, often investigating the antecedents of career patterns. Thus, the career perspective takes a longitudinal view, in contrast to the cross-sectional view taken by the occupational perspective. Vocational development is inferred and career is denoted by the systematic changes that can be observed in vocational behavior across two or more points in time. Because P-E fit can be studied at a single point in time, vocational behavior and its outcomes may be directly observable. In contrast, careers synthesize past memories, present reflections, and future anticipations into a narrative with a unified plot. As one's own story, a career is not directly observable. Accordingly, career denotes contemplation about the course of one's vocational behavior, not vocational behavior itself. This deliberation can focus on actual events such as one's occupations (objective occupational career) or on biographical themes (subjective vocational career) that give meaning to one's occupations.

The objective view of careers, whether by external observers or the person him- or herself, looks at the evolution of P-E fit across the life course. It is a précis of the actual events in one's work history, including all of the preparatory, participatory, and retirement positions an individual has occupied. This public definition of career can be objectively documented as facts and analyzed for patterns defined by the sequence and duration of positions occupied by an individual. The subjective view of careers as private meaning stands in contrast to the public pattern of occupations in a work history. Subjective career, while also a reflexive project, refers to how individuals understand their occupational roles and interpret their work experiences. A subjective career concentrates on the expression of a self-concept in work roles; it tells one's "own story," usually by focusing on a sense of purpose that coherently explains the continuity and change in one's self across time. It is this personal meaning that individuals use to orient themselves to their society's occupational structure, impose meaning on their vocational behavior, and plan their future work life. Whether one takes the public or the private view of career, most researchers who use the term career study the development of vocational behavior, or career development.

4. CAREER DEVELOPMENT

Careers involve the psychosocial integration of self-concepts and social roles. They are constructed by recursive transactions between individuals and their environments. Individuals, as self-organizing systems with their own goals, must find ways to fulfill their needs through playing social roles predetermined by the community in which they live. Communities usually script these social roles according to gender, race, ethnicity, and social class. Members of the community then socialize children by inducting them into these roles and communicating expectations for role performance. These social expectations are called developmental tasks, and failure to master the tasks can result in delayed or deviant patterns of adaptation. When the social expectations pertain to the work role, they are called vocational development tasks. Success in adapting to each task of vocational development results in more effective functioning as a student, worker, or retiree and lays the groundwork for mastering the next task along the developmental continuum. Skipping a task in the normal sequence usually results in difficulties in coping with later tasks. Super charted the linear progression of vocational development tasks and demarcated career stages that correspond to life stages of childhood, adolescence, early adulthood, adulthood, and late adulthood. The name of each career stage communicates its central activity: growth, exploration, establishment, maintenance, and disengagement.

4.1. Growth

The stage of career growth, usually ages 4 to 13, involves developing a vocational self-concept as well as the coping readiness and resources that will be needed to translate that self-concept into an occupational role. The vocational self-concept and

the translation attitudes and abilities constitute two interrelated yet distinct aspects of vocational development—its content and process. Career choice content centers on the vocational self-concept, that is, the totality of attributes that compose an individual's picture of self in the work role. Because vocational self-concepts develop primarily in social situations, an important determinant of their content is the social world that children inhabit. The social order, with its unequal arrangement of gender, race, and class, shapes children's occupational aspirations. Vocational self-concepts, and the careers they engender, are deeply grounded in an individual's internal representation of his or her location among unequal social positions. Using their "status identity," children circumscribe the range of occupational alternatives that they consider and later, when making choices, compromise their own interests to accommodate their perceptions of an occupation's status and sex type.

The process of circumscription and compromise in perceiving roles and aspiring to goals begins as children look to their parents as guides and identify with their preferences. This process of self-formation continues as children use imaginative play and school experiences to construct a bridge to take them from the family to the community. In what may be considered a very important career choice, children select role models who portray solutions to problems in growing up that they themselves face. As children imitate the desirable qualities of their role models, they rehearse relevant coping attitudes and actions, form values and interests in certain activities, and develop abilities and skills as they engage in these activities. Playing selective roles, with increasing attention to the results, enables a reality testing that strengthens or modifies vocational self-concepts.

While career choice content involves the subject matter of vocational self-concepts, career choice process deals with the attitudes and actions that construct a career. The construction process is initiated by the tasks of vocational development that individuals encounter during the growth stage of a career. Family, friends, and teachers expect students in middle school and junior high to develop a sense of concern about their vocational future, learn that they can control this future through self-determination and negotiation, become curious enough to explore possible selves and future scenarios, and develop confidence that they can implement plans and deal with problems that occur in actualizing choices. Attitudes of concern, control, curiosity, and confidence chart four developmental lines that, during adolescence, will coalesce to form a teenager's career choice readiness and resources.

4.2. Exploration

The career stage of exploration, usually ages 14 to 24, involves seeking information about self and about the work world in preparation for making choices that match the two. Three vocational development tasks characterize the years of career exploration: crystallizing vocational preferences, specifying an occupational choice, and actualizing that choice by obtaining a job in the specified occupation.

4.2.1. Crystallization

The first vocational development task of the exploration stage is to crystallize a vocational self-concept and translate that picture into preferences for a group of occupations, usually in similar occupational fields and at the same ability level. Crystallizing vocational preferences requires in-depth exploration to learn information about the self, occupations, and the fit between the two. As information and self-knowledge accumulate, individuals conceptualize new distinctions about their work-role self-attributes. Vocational exploration is fostered by educational experiences and can be accelerated by considering the results of vocational testing that draws an objective portrait of one's vocational interests, occupational abilities, and work values. A schema for understanding the world of work, such as the RIASEC hexagon, can be used to organize an individual's occupational information into a coherent framework that makes the information meaningful and relates it to vocational self-concepts.

4.2.2. Specification

Specifying an occupational choice from among crystallized vocational preferences is the second task of career exploration stage. Specification emerges from exploration in depth that sifts through attractive preferences and possible selves. This search involves more than matching vocational self-concepts to occupational outlets; it involves the psychosocial process of identity construction, because in declaring an occupational choice, individuals identify themselves before some audience. Stating an occupational choice constitutes a very public presentation of the self; it displays who we are and announces what we want to become. Public declaration of an occupational choice specifies

meanings that both an actor and observers attribute to the self as a worker. It involves telling a story about the self, one with a plot that makes a life whole by unifying and assigning meaning to various elements of the self.

A key step in crystallizing preferences and specifying choices involves developing the attitudes, beliefs, and competencies (the ABCs of career construction) that bear upon career decision making. The adaptive response dispositions and cognitive competencies have been identified and organized into a framework for assessing career adaptability. The dispositions are charted by the four developmental lines that originated in the growth stage and continue to develop during the exploration stage. Ideally, adolescents show career concern by orienting themselves to vocational development tasks and actively participating in coping with current tasks and preparing for anticipated tasks. This forward-looking attitude fosters the development of planning competency. Concern prompts the issue of control, as teenagers answer the question "Who owns my future?" Hopefully their answers reflect beliefs in self-determination and feelings of agency. These beliefs foster the development of competency in making matching decisions. Individuals who feel some sense of control over their future become curious about the educational and vocational choices they might make. There are many styles and strategies for career decision making, but the most adaptive ones seek information about possible selves and alternative choices. Curiosity about how to match oneself to occupational positions fosters increasing competence in self-knowledge and occupational information. The final disposition involves confidence that one can perform the behaviors needed to make career choices and to solve the problems that arise in implementing them. These feelings of career self-efficacy foster development of competence in problem solving.

From this view, the career choice process evolves over time and is facilitated by strategic and decisive attitudes toward making choices, focused by accurate conceptions of the decision-making process, moved by confidence about executing choice behaviors, and supported by competencies that inform choice behaviors. Of course, the development of career choice attitudes and aptitudes shows individual differences. For each individual, career choice readiness and resources develop at different rates, with possible fixations and regressions. Delays within or disequilibria between the four developmental lines produce problems in coping with the tasks of crystallization and specification. Moderate disharmony among the developmental lines produces variant patterns of indecision and unrealism, whereas strong disharmony produces deviant patterns of indecisiveness and indifference. Accordingly, the four developmental lines sketch a framework of career adaptability that can be used to assess career choice readiness and diagnose vocational decision-making difficulties, beginning in adolescence and continuing throughout adulthood.

Lack of readiness and resources for educational and vocational decision making lead to predictable career choice problems. These decisional difficulties have been studied since the 1930s. Originally, researchers investigated the construct of vocational indecision by assessing differences in personality between decided and undecided individuals. At first, it appeared that indecision was related to personality maladjustment; however, over time it became clear that delays in deciding might be developmentally appropriate and even adaptive. During the 1970s, researchers began to view "undecided versus decided" as a continuum rather than as two distinct statuses. Attention then turned to problems that delay or distort progressive movement along the continuum from undecided to decided.

Research has identified dozens of career choice problems and, through reflection and research, grouped these problems into homogeneous categories. Several alternative categorizations are available, yet each one depicts some combination of five types of career choice problems: informational, decisional, relational, motivational, and environmental. The most frequently experienced problems in making a career choice involve a lack of information about work-related self-attributes; about occupational requirements, routines and rewards; or about how to engage in the career choice process. The second most common group of career choice problems involves deficits in decision-making skill, including knowing and practicing the principles of decision making. Even when sufficient information is present, deficits in decisional competence may limit the ability to process information and make the decisions that produce a suitable occupational choice. The third group of career choice problems involves deficits in relational support, meaning that significant others such as parents or partners disapprove of a choice that attracts the individual or insist on a choice that repels the individual. The fourth category of career choice problems involves motivational problems that delay engagement in the choice process. These difficulties include dysfunctional career beliefs, anxiety about making choices, and adjustment

disorders. Whereas the fourth group of problems generally occurs before engaging in the choice process, the fifth group occurs after a choice has been made. Some individuals make choices that are difficult to convert into actuality because of internal or environmental barriers. The external barriers may involve inability to gain entry to or pay for training or apprenticeship programs as well as ageism, racism, and sexism. Internal barriers include lack of confidence and conflicting commitments.

It should be obvious that these career choice problems do not usually appear alone; typically, they occur in combinations and often form patterns. Therefore, researchers operationally define degree of indecision as the number of difficulties encountered, often distinguishing between undecided individuals who experience a few problems and indecisive individuals who experience many problems. Undecided individuals are generally considered to be experiencing a variant pattern of career development, with their inability to choose being viewed as developmentally appropriate and treated by information, decisional training, and social support. Indecisive individuals are generally considered to show a deviant pattern of vocational development in which chronic anxiety and lack of problem-solving skills characterize all their social interactions, not just those dealing with the work role. Indecisive individuals generally require counseling that concentrates on issues that cause their anxiety. Whether an individual finally chooses an occupation or lets time and circumstances dictate that choice, there comes a day when the individual must apply for a job.

4.2.3. Actualization

The third and final task of the exploration stage, actualizing an occupational choice, requires that individuals turn their verbal choice of an occupation into an actual fact. Actualizing a choice usually involves trying jobs in that occupation. Trial connotes employment in jobs that could become permanent, that is, adult employment as opposed to youth employment. The initial job allows individuals to try on the occupation for fit and then, through a process of elimination, subsequent trial jobs help the individual to zero in on a suitable job. The period of actualizing choices, which typically lasts about 4 to 7 years, is referred to as the school-to-work transition. During the years in which individuals seek to stabilize in a more-or-less permanent position, the critical vocational coping behaviors consist of actions that move them to increasingly more congruent jobs. Any action aimed at vocational movement counts, including job search activities such as reading job advertisements and applying for positions.

The six main types of vocational movement vary in quality, with three being adaptive (skilling, experimenting, and stabilizing) and three being maladaptive (drifting, floundering, and stagnating). Drifting involves random movement to a position for which the individual is no better suited than the position being left or for which the individual lacks aptitude, interest, or preparation. Drifting implies being carried along by others or wandering aimlessly from position to position. Floundering means stumbling performance in a position or plunging ahead from one unsuitable position to another. Stagnating involves staying longer than appropriate or remaining in a blind alley job; it involves deterioration because it hurts future occupational possibilities.

The more adaptive coping behaviors start with skilling, that is, preparing to enter a regular adult occupation through further schooling, training, and apprenticeship. Experimenting involves moving from one related job to another in a process that eventually zeros in on a suitable and viable position. Stabilizing results in eventually settling into and making a place for oneself in a more-or-less permanent position. Compared to experimenting, stabilizing shows a lack of doubt. By age 25, about 80% of people have stabilized in a position. Over the long-term career, the quality of vocational coping behavior during the school-to-work transition is more important that sticking with the first job. Adaptive individuals may change jobs many times, moving through progressively more congruent positions, until they stabilize in a suitable position. In contrast, less mature individuals may flounder, drift, and stagnate rather than take the risks involved in advanced trial and in-depth exploration.

4.3. Establishment

The years of career establishment, usually ages 25 to 44, involve making one's job secure. Establishment involves three developmental tasks that coincide with the early, middle, and late periods of young adulthood: stabilizing, consolidating, and advancing. These vocational development tasks are discussed first in terms of their direct outcomes, then are reviewed in terms of their indirect outcomes, namely, job satisfaction and organizational commitment.

4.3.1. Stabilizing

The major developmental task of the early period in the establishment stage requires that the individual stabilize in an occupational position, primarily by doing two things. First, the individual must adapt to the organizational culture that surrounds the job. Adaptation to the organization deals with participating in the work environment, not with performing job tasks. Organizational adaptation involves transactions between the worker and the job environment, with the person engaged in efforts to learn about the company of workers and the company engaged in efforts to socialize the newcomer. There can be some reciprocity in the newcomer changing the company, but this is infrequent, and minimal when it does it occur. Adaptive efforts can begin before the employee is hired. Organizational adaptation is smoother when one's expectations are met, so providing a job candidate with a "realistic job preview" is a useful practice that reduces surprises and confusion in newcomers.

As they begin their employment, greenhorns must learn how things are done in their new company. This includes learning about the people, politics, values, language, and history of the organization in general and of their work group in particular. It is important that novices learn how their job fits into the big picture, that is, why their jobs are important and what they contribute to the organization's goals. Organizational adaptation is furthered by learning about and complying with the company's formal policies and procedures as well as its informal practices and unofficial rules. Observing others do the job is important in "learning the ropes" and figuring out unstated rules. In addition to observational learning, prior work experience is helpful, as is taking advantage of orientation programs, buddy systems, and mentoring.

Position performance is the second task of the early period in establishment. In addition to adapting to the organizational culture, individuals must demonstrate competence at work. They must clearly understand their job tasks, take these responsibilities seriously, and exhibit efficient and effective work performance. This is accomplished by concentrating on the task at hand and doing it right the first time. One must also be flexible in learning new ways of doing the tasks. In addition to adhering to production standards and quotas, new employees must follow safety procedures that reduce accidents in different occupations. During idle time, workers in the early years of establishment must learn how to keep busy. If workers cannot perform the job tasks adequately, or cannot adapt to the company's culture, then instead of stabilizing they may leave voluntarily or be terminated. If they are comfortable doing the job, and doing it in that organization, then they can turn their attention to increasing their job security.

4.3.2. Consolidating

The tasks of the middle period of the establishment stage involve consolidating one's position for the long run. In contrast to position performance, which involves what a person does on the job, consolidation involves how one goes about doing that job. How one works has two components, one interpersonal and the other personal. The interpersonal component involves the vocational development task of forming congenial relationships with coworkers and supervisors. Coping behaviors that facilitate cooperation at work include communicating appropriately, supporting colleagues, encouraging coworkers, demonstrating goodwill to supervisors, tolerating individual differences, and dealing with diversity. Skill at organizational politics helps with this, as does finding safe ways to blow off steam. Of course, personality plays a significant role in coworker relationships, and lately attention has focused on how agreeableness and extroversion, two factors in the five-factor personality model (FFPM), may predict the enthusiasm and friendliness with which one engages coworkers. Inability to get along with coworkers is the most frequent reason for leaving or losing one's job during the middle years of establishment.

While the interpersonal component of how one works involves relationships on the job, the personal component of how one works involves work habits and attitudes. It is relatively easy to be industrious and dependable during a probationary period, or even during the first few years in a new position. A task of the middle years of establishment is to maintain positive attitudes and disciplined habits. Over time, some workers become complacent or lackadaisical in approaching their job tasks; they may even experience boredom. Of course, this posture diminishes their productivity and their value to the organization. Conversely, workers who meet the task usually increase their productivity as they expand their job knowledge and master the "tricks of the trade." Conscientiousness, as its very name indicates, is the personality factor in the FFPM that predicts success in maintaining adaptive work habits and realistic attitudes.

4.3.3. Advancing

After consolidating their positions, some workers, but not all, think about advancing to another position in their organization. To earn promotions, they must envision the career paths mapped by their organization and perform the behaviors needed to get ahead, which include showing initiative, taking on more responsibility, engaging in on-the-job training, and enrolling in continuing education. Workers who wish to advance, rather than maintain what they have established, yet see no way forward in their current organization may consider advancing their career by moving to a new organization that offers better opportunities. An even greater transition, or turn in the career path, involves advancement by preparing for and entering a new occupation. Changing occupations is best done by recycling through the tasks of exploration and establishment. The direct outcomes of mastering the vocational development tasks of stabilization, consolidation, and advancement are organizational adaptation, position performance, coworker relationships, work habits and attitudes, and promotions. The indirect outcomes are job satisfaction and organization commitment.

4.3.4. Job Satisfaction

Job satisfaction is the most frequently studied variable in organizational research because employers want to know how to avoid labor turnover and reward their best workers. In contrast to the organization's perspective on the relation of job satisfaction to productivity, vocational psychology views job satisfaction from the perspective of the worker, defining job satisfaction as the attitude or the general feeling that a worker has toward her or his job. The cognitive component of this attitude is an evaluative belief about how much one likes or dislikes a job. General job satisfaction can be explained by satisfaction with parts of a job that the worker likes or dislikes. When general satisfaction is high, little attention is paid to satisfaction with job facets. However, when job satisfaction in general is lower, examining the facets may explain the low morale by identifying the disliked job components. There are many facets, and many ways of conceptualizing and measuring them.

Vocational psychologists conceptualize job satisfaction as an indirect outcome of coping with the vocational development tasks of the establishment stage, so they link the facets of job satisfaction to the direct outcomes of coping with those tasks. When examining job dissatisfaction, they look first at the facet of P-E fit as played out in organizational adaptation and position performance. This includes variables such as satisfaction with the organization, including its policies, compensation and other benefits, as well as satisfaction with job duties, including the job's routines and requirements, work load, and empowerment. Next, they consider satisfaction with coworkers, communication patterns, supervisors, and then the employee's conscientiousness and work ethic. Finally, they consider satisfaction with opportunities for advancement. Satisfaction with advancement opportunities can produce career satisfaction, even when job satisfaction is low, if the worker views that job as a temporary rung on a meaningful career ladder. For example, a university instructor may be quite dissatisfied with her position and pay, yet be satisfied with her career progress, knowing that with her next promotion she will be an assistant professor, a goal she has held for 10 years.

In addition to person-position fit and adaptation to the developmental tasks of establishment, job satisfaction relates to two more variables. The first variable is psychological, meaning that it resides inside the person rather than in the person-position interaction. This psychological variable is negative affect. People who generally experience negative affect score lower on job satisfaction, as they do on measures of satisfaction with other life roles. The second variable, one from sociology, is life structure or the arrangement of an individual's social roles. For example, part-time work interacts differently than full-time employment with an individual's other social roles. Job dissatisfaction can be caused by, or be coincidental to, role conflicts and role overload.

Relationships between family and work have been frequently studied, with a particular concentration on the interaction between the work role and the family roles of spouse, parents, and child. Some workers can be dissatisfied with their work because the job does not permit flexible scheduling or the company is not family friendly. And, of course, there is spillover between domains, with a divorce at home deteriorating work performance or a demotion at work causing arguments at home. Relationships between leisure and work have also drawn significant study. For example, type A personalities may experience high job satisfaction yet low life satisfaction if time and behavior at work reduce their engagement in other core roles such as recreation and friendship. Conversely, job dissatisfaction can be ameliorated by leisure, say by enjoying on a fishing line that which cannot be enjoyed on an assembly line.

The effects of job satisfaction are addressed by two relatively distinct research literatures. One deals with leaving dissatisfying jobs, while the other deals with being forced to stay in unsatisfying positions. Workers who must remain in unsatisfying positions may experience job stress that reduces their general emotional well-being. When people are chronically dissatisfied with a job they must perform 40 hours a week, they may eventually suffer from job burnout, a condition associated with a wide range of physical and psychological symptoms—headache, upset stomach, anxiety, and depression. Job dissatisfaction has even been linked to life expectancy, with those who hate their work predicted to have a shorter life expectancy.

Of course, an alternative to suffering the emotional and physical consequences of chronic job dissatisfaction is job withdrawal, including absenteeism and quitting. Although job dissatisfaction is widely believed to relate to absence, the empirical support is weak because there are many other reasons for being absent, including illness, personal business, and family responsibilities. In fact, having primary responsibility for the care of a child predicts absence better than does job dissatisfaction. Compared to absenteeism, there is stronger evidence of a causal link between job dissatisfaction and turnover. Dissatisfaction may lead to an intention to leave, followed by engagement in job search activities that include writing a resume, enrolling with job-finding agencies, and interviewing for positions. If another job is found, then the dissatisfied employee can leave. However, without a job offer in hand, employees are unlikely to quit. Also, it must be noted that in deciding whether to resign a position, individuals think about more than just how they feel about their jobs. They also think about the interpersonal relationships that they may have established at work.

4.3.5. Organizational Commitment

When work relationships are characterized by social support and fair treatment, individuals may feel an attachment to the organization and prefer not to leave their colleagues. This attachment between a worker and others in the work environment is called commitment. Typically researchers concentrate on organizational commitment in general, yet commitment can be connected to other foci such as coworkers, supervisors, the job, and the occupation. In addition to multiple foci, commitment itself is multidimensional, typically characterized by three dimensions. The affective dimension of commitment denotes emotional attachment and identification with the organization, job, or coworkers. These bonds make one want to remain in the organization. The continuance dimension of commitment denotes the need to stay because no other job is available or, more often, because quitting would cost too much or require too many sacrifices. The normative dimension of commitment denotes a perceived obligation to stay because leaving would be wrong or let down too many people. Affective and normative commitment correlate positively to attendance, productivity, organizational citizenship behaviors, and pleasant affective experiences at work. Continuance commitment is typically unrelated, or even negatively correlated, to positive work behaviors. Of course, the affective, continuance, and normative dimensions of commitment can interact. For example, low affective commitment with high continuance commitment can cause problems such as making an individual feel locked into a situation with which they do not identify.

The normative dimension of commitment brings to mind the "psychological contract" between the employee and the organization. This construct refers to beliefs about reciprocal obligations held by people in an exchange relationship, or simply stated, issues of loyalty and payback. A psychological contract involves an employee's perception of explicit and implicit promises regarding the exchange of the worker's effort, abilities, and loyalty for the organization's pay, promotions, and security. Individual workers prefer clear, unambiguous promises because clarity makes it easier to evaluate promises made and promises kept. Fulfilled contracts increase job satisfaction and commitment, whereas breached contracts lead to dissatisfaction and withdrawal behaviors. Breaches in the psychological contract require justification; these appraisals of workplace justice concentrate on fairness and inclusiveness in the process and outcome of reward allocation.

4.4. Maintenance

At some point in a career, whether it be successful or not, the worker seriously evaluates issues of satisfaction and commitment. During this period of renewal, individuals engage in self-questioning. They ask themselves, sometimes prompted by a partner, the midlife question: "Do I want to do this for the next 25 years?" Essentially, they engage in a process of re-evaluating their identity as they ask themselves, and sometimes their family and friends, if they should hold on or let go. If they decide to change organizations or occupations, then they recycle through

the tasks of exploration and establishment by crystallizing and specifying a different choice and then actualizing and stabilizing a new position. What we have been considering as a maxicycle of career stages can, in terms of a position change, be considered a minicycle of job transition. Having stabilized and maybe even advanced in a position, an individual may experience new growth that raises identity questions and leads to advanced exploration and then establishment in new position. Career patterns are defined by the repetition, frequency, and duration of these minicycles.

In the large, bureaucratic organizations of the 20th century, individuals might go through the maxicycle and maintain themselves in the stabilized position for 30 years before retiring. At the other extreme are individuals who frequently traverse the transition minicycle as they repeatedly grow, explore, stabilize, and then move to positions of new growth. Between these two extremes of a job for life and a job for now lie most of today's workers, who every 5 or 10 years find their career destabilized by illness and injury, layoffs and downsizing, or social change and personal upheaval. Developing a career today can be more about management of one's skills and opportunities than maintaining oneself in an established position.

When people do maintain themselves in a job, they enter the maintenance stage of a career, generally defined as ages 45 to 64, or from midlife to retirement. Maintenance addresses the requirement to remain in place, take care of what one has produced, and continue to produce. Society expects mature adults to hold steady in the positions they occupy, actively strive to preserve their status, remain interested in their work, and stay committed to their organizations. To remain in place, one must cope with novel, non-maturational problems that arise from environmental change or inner experience and preserve one's self-concept through modifying and adjusting one's views and beliefs. In contrast to the prior career stages, the maintenance stage is not characterized by tasks that progress in a predictable order related to age-related social expectations; the focus is on preservation, not progress. Consequently, research on this career stage has concentrated on the styles with which individuals retain their positions and preserve their identities.

Three basic styles of maintenance have been identified: upholding, updating, and innovating. The first style, upholding, involves continuing to do what one has done before, maintaining the same standards, and keeping things going in good condition. Workers who uphold their positions can hold their head up in pride about what they continue to do. They take care of what they have accomplished and continue to be industrious. In contrast to this active upholding, there is passive holding on, in which one is more careless at work and shifts attention to waiting for retirement or looking for an early exit from the workforce. Rather than being careful in what they do, workers who hold on care more about how they are doing than what they are doing. They are careful how they behave when supervisors are watching. Instead of holding up their heads with pride, they hold up the organization by not doing their tasks as they should be done and in extracting rewards they do not deserve. The second positive style of maintenance, after upholding, is doing tasks better by updating. Workers using this style of maintenance adapt efficiently and respond to changing technology and tasks by keeping up with new developments. They stay fresh by updating their skills and knowledge, and in the process they reaffirm their self-concepts. While updating means doing tasks better, the third positive style of maintenance is to do tasks differently. Innovating involves actively seeking new ways of doing routine tasks and discovering new challenges. Of course, as they break new ground, some innovators may eventually decide to recycle to new positions and re-establish themselves. Each style, whether it be upholding, updating, or innovating, involves different trade-offs, yet workers using each style can maintain job success, satisfaction, and stability as well as provide security for their families.

4.5. Disengagement

At some point in a career, energy wanes and retirement looks good. It is time to decelerate by turning tasks over to others and planning for retirement. The task of actually disengaging from the work role goes better if the worker can decelerate first. Similar to other major transitions, small steps smooth the process. Deceleration can be accomplished by phase-outs, in which the individual works fewer and fewer hours over a period of time, occasionally lasting as long as several years. Another form of deceleration is called "bridge employment," in which an individual retires from a permanent position and then, before complete retirement, works at a part-time job, becomes self-employed, or works full-time at a temporary job. Bridge employment is becoming increasingly common as four out of five retirements are now early retirements and older workers remain in excellent health

well into their 70s, a trend that is compounding age discrimination in employment.

Deceleration, whether with one's permanent employer or through bridge employment, enables individuals to avoid the swift social isolation, reduced income, and bereavement that can accompany retirement from the work role. Nevertheless, at some point, the worker who lives long enough finally must retire and adjust to retirement living. Often, retirees engage in hobbies they once enjoyed during late childhood, bringing the career cycle full circle. If they live near grandchildren, they can once again play the games they enjoyed as children, thereby helping the next generation of workers begin to grow their careers and prepare to sustain our world. For, as Freud intimated, the world we know would perish without work.

See Also the Following Articles

Career Counseling ■ Holland's Theory (Vocational Personality Types) ■ Indecision, Vocational ■ Job Analysis, Design, and Evaluation ■ Organizations, Careers in ■ Person–Environment Fit ■ Traits ■ Vocational Assessment in Schools ■ Vocational Interests ■ Work Adjustment ■ Work Environments

Further Reading

Holland, J. L. (1997). *Making vocational choices: A theory of vocational personalities and work environments* (3rd ed.). Odessa, FL: Psychological Assessment Resources.

Kelly, K. R., & Lee, W. (2002). Mapping the domain of career decision problems. *Journal of Vocational Behavior*, *61*, 302–326.

Parsons, F. (1909). *Choosing a vocation*. New York: Agathon Press.

Savickas, M. L. (2002). Career construction: A developmental theory of vocational behavior. In D. Brown and Associates (Eds.), *Career choice and development* (4th ed., pp. 149–205). San Francisco, CA: Jossey Bass.

Savickas, M. L., & Spokane, A. R. (Eds.) (1999). *Vocational interests: Their meaning, measurement and counseling use*. Palo Alto, CA: Davies-Black.

Strong, E. K., Jr. (1943). *The vocational interests of men and women*. Palo Alto, CA: Stanford University Press.

Sterns, H. L., & Huyck, M. H. (2001). The role of work in midlife. In M. Lachman (Ed.), *Handbook of midlife development* (pp. 447–486). New York: Wiley.

Super, D. E. (1984). Career and life development. In D. Brown, & L. Brooks (Eds.), *Career choice and development* (pp. 192–234). San Francisco, CA: Jossey-Bass.

Walsh, W. B., & Savickas, M. L. (Eds.) (in press). *Handbook of vocational psychology: Theory, research, and practice* (3rd ed). Hillsdale, NJ: Erlbaum.

Well-Being

Richard E. Lucas
Michigan State University, East Lansing, Michigan, USA

Ed Diener
University of Illinois at Urbana-Champaign, Champaign, Illinois, USA

1. The Hierarchical Structure of Subjective Well-Being
2. Measurement Issues
3. Theoretical Issues
4. Effects of High Well-Being
5. Summary
   Further Reading

GLOSSARY

affective dimension of well-being Emotions and moods.
emotions Distinct evaluative (i.e., positive or negative) reactions to specific events.
experience sampling method A tool for assessing a person's subjective well-being that requires participants to track their well-being over time.
life domains Separable aspects of a life, including health, family life, occupation, social life, etc.
life satisfaction An individual's judgment that his or her life is going well.
moods Diffuse, affective feelings that are not necessarily tied to specific events.
negative affect Emotions and moods such as fear, anger, and sadness.
positive affect Emotions and moods such as happiness, joy, and excitement.
subjective well-being Well-being from an individual's own perspective.

Subjective well-being (SWB) is a broad construct that comprises multiple, separable components. These components are linked by a few key characteristics. First, as its name suggests, subjective well-being refers to well-being from an individual's own perspective. Although external observers may agree that the conditions in a person's life are less than ideal, that person may still feel a sense of satisfaction and well-being. SWB researchers recognize that the effects of objective circumstances are filtered through individuals' unique interpretations and goals, as well as their current life stage. Thus, the field of SWB places much weight on subjective evaluations. Second, SWB includes both positive and negative aspects of an individual's life. Although SWB researchers are interested in the factors that might prevent or improve negative states such as anxiety and depression, they are also interested in moving people from neutral to more positive states. Thus, positive constructs such as life satisfaction and positive affect play an important role in SWB research. Finally, SWB researchers often focus on evaluations of an individual's life as a whole. Although this can be accomplished either by asking for global judgments or by aggregating across numerous more specific judgments (including specific domain satisfactions or momentary experiences), SWB researchers often strive to understand how various circumstances and characteristics combine to create a life that is considered to be positive. Thus, SWB researchers value the utilitarian goal of creating the

©2004 Elsevier Inc.
All rights reserved.

greatest happiness for the greatest number of people, and they believe that by studying the causes, correlates, and outcomes of high levels of well-being, this goal can be attained. SWB constructs can be useful in a number of applied settings. First, because high well-being is a valued goal for most people, well-being can often serve as a useful outcome measure. Psychologists interested in improving work or community environments could measure SWB to determine whether interventions are successful. In addition, modern SWB research focuses on the effects that SWB can have on other outcome variables. It is possible that by measuring and even changing levels of SWB, psychologists will be better able to predict and affect important practical outcomes. This article reviews the topics that will be most useful to applied psychologists interested in understanding well-being. First, an overview of the components that are typically linked under the overarching construct of SWB are discussed. Then, the techniques that can be used to measure these components are presented. Finally, theoretical issues regarding the causes and consequences of SWB are discussed.

1. THE HIERARCHICAL STRUCTURE OF SUBJECTIVE WELL-BEING

Diener *et al*. noted that the construct of subjective well-being (SWB) has a hierarchical structure. At the highest and most abstract level is an overall evaluation of a person's life. Although the goal of SWB researchers is to understand this high-level evaluation, such an evaluation is difficult to measure directly. Instead, psychologists usually assess a variety of lower level constructs. For instance, researchers often distinguish between the affective and cognitive dimensions of well-being. On the affective side, researchers focus on emotions and moods. Emotions are thought of as distinct evaluative (i.e., positive or negative) reactions to specific events, whereas moods are often described as diffuse, affective feelings that are not necessarily tied to specific events. Presumably, if one's life is going well, he or she will experience high levels of positive emotions and moods and low levels of negative emotions and moods.

There are both theoretical and empirical reasons to distinguish positive affect (which includes such emotions and moods as happiness, joy, and excitement) from negative affect (which includes such emotions and moods as fear, anger, and sadness). Distinct emotions often correlate more strongly with other emotions of the same valence than they do with emotions of the opposite valence. In addition, positive and negative affect are usually not so strongly correlated as to suggest that they reflect opposite poles of the same underlying dimension (although there is still debate about the bipolarity of affect). Furthermore, these separable aspects of affective experience are often differentially related to distinct predictor and outcome variables. For instance, positive affect is more strongly related to the personality trait of extraversion than is negative affect, whereas negative affect is more strongly related to the personality trait of neuroticism than is positive affect. Therefore, it is important to assess these variables separately.

Both positive and negative affect can also be distinguished from the more cognitive component of life satisfaction. Life satisfaction reflects an individual's judgment that his or her life is going well. It may be possible, for instance, for an individual to experience high levels of negative affect and still be satisfied with the conditions in his or her life. He or she may not value emotional experiences or may be pursuing some goal that is inherently satisfying yet which often leads to frequent experiences of negative affect. In 1996, Lucas *et al*. used multitrait-multimethod techniques to show that life satisfaction, positive affect, and negative affect were related but discriminable constructs.

The fourth component of SWB is domain satisfactions. Presumably, when constructing a life satisfaction judgment, an individual examines the various domains in his or her life (including health, family life, occupation, and social life), weighs the importance of these domains, and then aggregates the various judgments to get an overall evaluation of satisfaction with life. SWB researchers know, however, that this process does not occur perfectly. People do not have the motivation or ability to compile and aggregate across an exhaustive list of life domains. Therefore, when trying to understand the ways that people make life satisfaction judgments, or simply to ensure that certain important life domains are included in an overall evaluation, psychologists may want to assess satisfaction with specific life domains.

The hierarchical structure of SWB does not end with the four components of positive affect, negative affect, life satisfaction, and domain satisfactions. Each can be further divided. For instance, some researchers may be interested in specific emotions or reactions to specific emotion-eliciting experiences. Similarly, researchers may be interested in both satisfaction with a broad domain of life and satisfaction with narrower features

of that domain. Thus, for certain purposes, it may be useful to develop and use more precise measures of constructs that fall at a lower level. However, for most psychologists, the four components will provide a broad overview of an individual's subjective evaluation of his or her life.

2. MEASUREMENT ISSUES

Although the fact that SWB is a subjective phenomenon necessitates the use of some self-report methods, these are not the only measurement techniques that SWB researchers use. Instead, most SWB researchers share the view that multimethod assessment will provide the most informative perspective on the subjective quality of a person's life. In fact, by comparing information attained through different methods, SWB researchers hope to understand how evaluations are made, which in turn affects understanding of the factors that influence the construct. Next, the various methods that have been used to assess SWB are reviewed.

2.1. Self-Reports

The simplest, most cost-effective, and most intuitive way to assess SWB is simply to ask individuals how they would evaluate their lives. There are many different self-report measures from which to choose. For instance, self-report measures can include single items, such as "All things considered, how satisfied are you with your life?" or "In general, how happy are you?" Although single-item measures can be problematic (they may have poor reliability and random error cannot be canceled by aggregating across multiple items), existing research using these measures suggests that they have some degree of reliability and validity. For instance, in 2002 Diener *et al*. showed that a single-item measure of cheerfulness correlated very strongly with an aggregated multiple-item positive affect scale that had been administered multiple times during a 4-month period. More impressively, this single-item cheerfulness measure was able to predict respondents' income 20 years later. Lucas *et al*. showed that a single-item life satisfaction scale was sensitive to life changes such as marriage, widowhood, and unemployment. In addition, this measure also prospectively predicted the experience of life events. Thus, single-item measures can be useful when longer multiple-item scales cannot be used.

Although single-item scales show promise, multiple-item scales are desirable when feasible. Multiple-item scales are likely to be more reliable, they allow reliability to be assessed, and they often capture a broader construct than a single-item is able to assess. There are multiple-item measures that tap each of the components of SWB described previously. For instance, in 1985 Diener *et al*. created a five-item scale to measure life satisfaction. This scale correlates strongly with other SWB measures and can distinguish among various groups that are expected to differ in levels of well-being. Lucas, Diener, and Larsen and Larsen, Diener, and Lucas review a number of measures that can be used to assess positive and negative affect. Most domain satisfaction scales simply ask respondents to indicate how satisfied they are with a variety of areas in their lives.

2.2. The Experience Sampling Method

An important variation on standard self-report assessments is the experience sampling method (ESM). Unlike global questionnaire measures, ESM assessments require participants to track their well-being over time. For instance, participants may carry handheld computers that signal alarms multiple times per day. Each time an alarm sounds, the participant completes a brief questionnaire. This technique allows researchers to avoid certain memory and aggregation problems that occur when respondents are asked to make global judgments, and it also allows researchers to investigate more dynamic aspects of well-being. For instance, psychologists can investigate the extent to which well-being fluctuates over time or correlates with various situational features.

2.3. Problems with Self-Reports

As with any self-report measure, self-reports of well-being are not without their problems. These reports have been shown to be influenced by a variety of irrelevant situational factors, including current mood, current weather, question wording, and availability of response options (although some studies have suggested that the effect of these factors is not strong). Thus, additional methods of assessment are useful when it is feasible to obtain them. One common alternative method is the use of observer reports.

2.4. Observer Reports

Although individuals are thought to have privileged information about their own subjective feelings of well-being, signs of these subjective feelings may be exhibited to others. For instance, individuals may share their thoughts about subjective well-being with close friends and family members, or these close acquaintances may be able to interpret nonverbal signs of well-being, such as smiling and crying. Thus, researchers interested in using multimethod assessment often obtain observer reports of well-being constructs. At their simplest, observer reports may consist of standard self-report instruments in which the wording has been changed to reflect the fact that respondents are rating someone other than themselves. In general, such informant reports correlate moderately with self-reports of well-being. It is also possible to obtain observer reports from individuals who do not know the target well. For instance, expert raters may be trained to recognize specific indicators of emotional experience, including nonverbal behaviors or even specific facial muscle movements.

2.5. Problems with Observer Reports

Although observer reports may prevent certain biases associated with self-report from occurring, they have their own set of problems. Observers cannot see subjective feelings directly, and they must infer subjective states from what the target person says and the behavior that the target person exhibits. Thus, both self-reports and informant reports can be supplemented with additional measurement techniques.

2.6. Physiological Indicators

Recent advances in psychophysiology and neuroscience have allowed researchers to identify specific indicators that tend to be linked with emotional experience. For instance, heart rate, blood pressure, bodily temperature, and skin conductance have been shown to be related to certain types of emotional experience. In addition, certain patterns of brain activity can be assessed using electroencephalograms, positron emission tomography scans, and magnetic resonance imagery. These patterns have been shown to correlate with individual differences and within-person changes in emotions.

2.7. Nonphysiological Indicators

Other researchers look to nonphysiological indicators of emotions and well-being. For instance, in 1998 Rusting reviewed evidence that trait and state levels of affect are associated with distinct cognitive processes that can be reliably assessed. For example, happy people are often quicker to recognize positive words than are unhappy people, and happy people are more likely to come up with positive words than negative words in word-completion tasks. Both cognitive and physiological measures are likely to correlate only weakly with self-reports of well-being. Thus, at this point, they probably cannot be entirely substituted for self-reports. However, non-self-report techniques can be a useful addition to a multimethod study, particularly when there are concerns about social desirability or other response problems.

3. THEORETICAL ISSUES

Psychologists who are interested in developing theories of well-being generally focus on two questions: What are the factors that lead to high levels of well-being? and What effects do high levels of well-being have in people's lives? Whereas the first question guided early research in the field, modern research increasingly focuses on the latter. Thousands of studies have been conducted, and researchers now know a great deal about these questions.

3.1. Demographic and Situational Characteristics

Early research on subjective well-being tended to focus on the demographic and situational characteristics that correlate with well-being. Perhaps the most surprising finding from these studies was that external circumstances seem to account for very little variance in reports of well-being. For instance, when researchers assess a wide variety of demographic factors (sex, age, income, health, etc.) and life circumstances, these variables only account for 15–20% of the variance in well-being measures. Even when extreme circumstances are examined, the effects are not large. In perhaps the most famous of these studies, Brickman *et al.* in 1978 showed that individuals who had won large sums of money in lotteries were not much happier than a matched control group. Similarly, individuals with spinal cord injuries were not much less happy than a

matched control group. One possible explanation for this finding is that individuals initially react to life events but then gradually adapt back to their previous levels. Support for this hypothesis was found in longitudinal studies that traced SWB after the experience of these life events. For instance, in 1996 Suh *et al*. found that only recent life events (those occurring within the past 3 months) affected current levels of well-being.

3.2. Adaptation Processes

However, recent large-sample longitudinal studies have suggested that adaptation processes are complicated and that life events can have important long-term consequences for people's well-being. Lucas *et al*., for instance, showed that the experience of unemployment causes long-term changes in people's levels of life satisfaction. Although people who had been unemployed in the past started to rebound toward their initial levels of satisfaction following a bout of unemployment, they seemed to be permanently affected by this experience and tended not to return all the way back to initial levels. Similarly, although individuals who experienced the death of a spouse did return very close to their prewidowhood levels of satisfaction, this process was very slow, taking an average of 8 years. Furthermore, Lucas *et al*. showed that there is much variability in people's reactions to many different types of life events. Although, on average, people may adapt to events such as marriage, there are considerable individual differences in reaction and adaptation, and some people experience long-term changes in satisfaction following these events.

3.3. Traits and Personality Characteristics

These individual differences in reactions to life circumstance suggest that certain personality characteristics play an important role in determining levels of SWB. In contrast to the weak effects of demographic characteristics, personality traits are often moderately to strongly correlated with well-being. In particular, the personality traits of extraversion and neuroticism have been identified as moderate to strong correlates of SWB. Although it is unclear exactly why these constructs are related, a number of researchers have linked personality and affect through underlying physiological systems. For instance, in 1981 Gray suggested that there are two physiological systems that underlie the personality traits of extraversion and neuroticism: a behavioral activation system (BAS) that governs reactions to signals of reward and nonpunishment and a behavioral inhibition system (BIS) that governs reactions to punishment and nonreward. Tellegen and others have suggested that the BAS is closely linked to extraversion and positive emotionality, whereas the BIS is closely linked to neuroticism and unpleasant affect. The suggestion that extraversion and neuroticism share common physiological underpinnings with positive and negative affect is supported by the fact that affective traits are moderately stable over time and have moderate to high heritability.

3.4. Nontrait Personality Characteristics

It is also likely that additional nontrait personality characteristics can directly affect well-being or can serve to moderate the effects of external circumstances on well-being. For instance, in 1986 Emmons suggested that by understanding individuals' goals, SWB researchers can better predict well-being outcomes. Emmons showed that simply having valued goals is associated with higher levels of well-being. In addition, various characteristics of one's goals (including whether goals conflicted with one another) are related to SWB outcomes. Other researchers have shown that the effect of specific external circumstances and demographic characteristics such as one's income may differentially affect well-being depending on one's goals and values.

3.5. Circumstances in Context

Finally, it is important to consider life circumstances in the context of the lives in which they are being experienced. In 1976, Campbell *et al*. argued that these life circumstances only have meaning when they are compared to various standards that an individual might have. For instance, a certain income level may be seen as desirable or undesirable depending on the income that an individual had in the past, his or her expectations regarding that income, or the income that salient and proximal individuals are earning. Thus, it may be important to assess various contextual factors and standards of comparison before a complete understanding of the factors that influence well-being can be attained. However, in 1997 Diener and Fujita reviewed evidence showing that even these comparison

processes are flexible and function differently across individuals. For instance, it is not the case that surrounding oneself with individuals who are worse off than oneself will inevitably lead to higher well-being. Although it is possible that individuals will recognize that they are better off than those with whom they can compare, this comparison information can be used in different ways. On the one hand, downward comparisons can make one value one's current status, which could lead to high well-being. On the other hand, downward comparisons can remind individuals that things can be much worse. This pessimistic interpretation might lead to lower well-being. The choice of comparison targets and the way this comparison information is used vary across individuals.

4. EFFECTS OF HIGH WELL-BEING

Although fostering high levels of SWB is itself an important and valued utilitarian goal, other researchers have wondered about the effects of high well-being. Is it the case that happy people become apathetic and lazy because they feel like they have achieved all their goals? Or do happiness, positive affect, and life satisfaction serve to motivate individuals to reach even higher goals? Much recent research seeks to answer these questions. If SWB is functional, then this will have important implications for applied psychologists.

4.1. Happiness and Productivity

There is a long line of research in the organizational literature that is dedicated to determining whether happy workers are more productive than their less happy coworkers. Much of the early research on the happy/productive worker hypothesis was based on the very simple idea that well-being (particularly the job satisfaction component on which most early research focused) was an attitude. If attitudes affect behavior, then positive attitudes should elicit positive behaviors. Recently, well-being researchers have suggested that affect and satisfaction may function in many more complex ways, and these functions may have implications for outcomes such as worker productivity.

4.2. Positive Social Outcomes

Based on a 2003 review by Lyubomirsky *et al.*, Lucas and Diener specified six types of outcomes that may be relevant for the happy/productive worker hypothesis. First, there is a great deal of evidence that SWB, particularly the positive affect component, is linked with positive social outcomes. Correlational research shows that feelings of sociability and social activity are correlated with positive affect. In addition, both correlational and experimental studies show that induced positive mood can cause changes in helpfulness, cooperation, and friendliness and self-disclosure. Happy people like other people more than do unhappy people, and people like happy people more than unhappy people. Thus, higher levels of positive affect are likely to lead to greater social outcomes and increases in one's social resources.

Happiness may also lead to benefits in motivation. For instance, researchers interested in the structure of mood have shown that feelings of energy and activity are strong markers of the positive affect dimension. Correlational experience sampling studies have also shown that people feel and act more energetic and active when they are feeling happy. In 1988, Cunningham experimentally manipulated positive affect and showed that when in positive moods people were more likely to prefer active and exciting activities. Thus, happy people may simply have more energy to complete important tasks.

Perhaps more important, however, there appear to be effects beyond just boosts in energy. Happy people are also more likely to set higher goals and approach those goals with greater confidence. Correlational evidence shows that happy individuals are more confident and assertive, and experimental evidence shows that manipulating happiness leads to higher levels of self-confidence and more positive self-evaluations. Furthermore, manipulating levels of affect can lead to changes in the goals that individuals set for themselves. Thus, if happy individuals set higher goals and have greater confidence about their potential for success, they may, in fact, achieve more than their less happy counterparts.

A fourth potential benefit of subjective well-being is that a variety of emotional variables are empirically linked with health and physical well-being. Specifically, individuals who experience high levels of stress and negative affect experience more health problems than do individuals with low levels. In addition, there is increasing evidence from both correlational and experimental studies that positive mood can also have a protective effect, although it is unclear whether these effects are direct or whether they simply moderate the effects of negative affect on health. These protective effects may even translate into greater longevity for happy individuals.

4.3. Effects of Happiness and Well-Being in the Cognitive Realm

Finally, there are two potential effects of happiness and well-being in the cognitive realm. First, there is evidence from experimental studies of induced mood that higher levels of happiness and positive affect lead to greater creativity. Individuals who have undergone positive mood inductions tend to make more unusual responses to word association tasks, to see more creative connections among concepts, and to list more unusual exemplars for categories. In situations in which novel and creative responses are required, individuals who are in a positive mood may have an advantage.

In addition, mood seems to affect the specific cognitive strategies that individuals are likely to use. On the one hand, unhappy individuals have often been shown to think more carefully about arguments and to rely on stereotypes less heavily than do happy people. This may lead to an advantage for less happy people when attention to detail is required. Happy people, on the other hand, may process information more efficiently, using effective heuristics. Each of these strategies may be more or less useful depending on the specific situation. At this point, it is also unclear whether distinct emotional states lead directly to the differential use of cognitive strategies. Martin *et al*., for instance, in 1993 argued that moods simply provide information about the state of the world. This information is then used as one factor in the decision to use a specific strategy. Martin *et al*. suggested that mood is not inevitably linked with a particular cognitive outcome. In support of this viewpoint, in 1994 Bodenhausen *et al*. found that although individuals in positive moods were more likely than less happy individuals to use stereotypes when making judgments, the happy individuals could overcome these stereotypes if they were told they would be held accountable for their decisions.

5. SUMMARY

Applied psychologists will find SWB constructs to be useful in a variety of settings. On the one hand, SWB can be seen as the ultimate utilitarian outcome measure. Presumably, any intervention that is designed to effect change in an individual, an organization, or a community could be evaluated in terms of its effect on SWB. In addition, SWB may be functional. Various components of well-being have been shown to have a causal impact on such factors as sociability, motivation, health, and cognition. Therefore, intervention programs aimed at increasing well-being may, in turn, lead to changes in many other areas. To successfully incorporate SWB measures into their programs, applied psychologists must first determine which aspect of SWB they wish to study and then develop measures that can adequately tap these dimensions.

See Also the Following Articles

Emotion ■ Social Comparison and Subjective Well-Being ■ Traits

Further Reading

Bolger, N., Davis, A., & Rafaeli, E. (2003). Diary methods: Capturing life as it is lived. *Annual Review of Psychology*, *54*, 579–616.

Cacioppo, J. T., Berntson, G. G., Larsen, J. T., Poehlmann, K. M., & Ito, T. A. (2000). The psychophysiology of emotion. In M. Lewis, & J. M. Haviland-Jones (Eds.), *Handbook of emotions* (2nd ed., pp. 173–191). New York: Guilford.

Danner, D. D., Snowdon, D. A., & Friesen, W. V. (2001). Positive emotions in early life and longevity: Findings from the nun study. *Journal of Personality Psychology*, *80*, 804–813.

Diener, E., Nickerson, C., Lucas, R. E., & Sandvik, E. (2002). Dispositional affect and job outcomes. *Social Indicators Research*, *59*(3), 229–259.

Diener, E., & Oishi, S. (2000). Money and happiness: Income and subjective well being across nations. In E. Diener, & E. M. Suh (Eds.), *Culture and subjective well being* (pp. 185–218). Cambridge, MA: MIT Press.

Diener, E., Scollon, C. N., & Lucas, R. E. (2004). The evolving concept of subjective well being: The multifaceted nature of happiness. *Advances in Cell Aging and Gerontology*, *15*, 187–219.

Eid, M., & Diener, E. (in press). Global judgments of subjective well being: Situational variability and long term stability. *Social Indicators Research*.

Harker, L., & Keltner, D. (2001). Expressions of positive emotions in women's college yearbook pictures and their relationship to personality and life outcomes across adulthood. *Journal of Personality and Social Psychology*, *80*, 112–124.

Judge, T. A., Thoresen, C. J., Bono, J. E., & Patton, G. K. (2001). The job-satisfaction–job performance relationship: A qualitative and quantitative review. *Psychological Bulletin*, *127*, 376–407.

Larsen, R. J., Diener, E., & Lucas, R. E. (2002). Emotion: Models, measures, and individual differences. In

R. G. Lord, R. J. Klimoski, & R. Kanfer (Eds.), *Emotions in the workplace* (pp. 64–106). San Francisco: Jossey-Bass.

Lucas, R. E., Clark, A. E., Georgellis, Y., & Diener, E. (2003). Reexamining adaptation and the set point model of happiness: Reactions to changes in marital status. *Journal of Personality and Social Psychology*, *84*(3), 527–539.

Lucas, R. E., Clark, A. E., Georgellis, Y., & Diener, E. (2004). Unemployment alters the setpoint for life satisfaction. *Psychological Science*, *15*, 8–13.

Lucas, R. E., & Diener, E. (2003). The happy worker: Hypotheses about the role of positive affect in worker productivity. In M. R. Barrick, & A. M. Ryan (Eds.), *Personality and work* (pp. 30–59). San Francisco: Jossey-Bass.

Lucas, R. E., Diener, E., & Larsen, R. J. (2003). Measuring positive emotions. In S. J. Lopez (Ed.), *Positive psychological assessment: A handbook of models and measures* (pp. 201–218). Washington, DC: American Psychological Association.

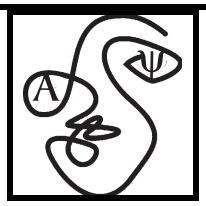
Meloy, M. G. (2000). Mood-driven distortion of produce information. *Journal of Consumer Research*, *27*, 345–359.

Salovey, P., Rothman, A. J., Detweiler, J. B., & Steward, W. T. (2000). Emotional states and physical health. *American Psychologist*, *55*, 110–121.

Watson, D., Wiese, D., Vaidya, J., & Tellegen, A. (1999). The two general activation systems of affect: Structural findings, evolutionary considerations, and psychobiological evidence. *Journal of Personality and Social Psychology*, *76*(5), 820–838.

Wright, T. A., & Staw, B. M. (1999). Affect and favorable work outcomes: Two longitudinal tests of the happy-productive worker thesis. *Journal of Organizational Behavior*, *20*, 1–23.

Zelenski, J. M., & Larsen, R. J. (2000). The distribution of basic emotions in everyday life: A state and trait perspective from Experience Sampling data. *Journal of Research in Personality*, *34*(2), 178–197.



Women's Health

Pilar Sánchez-López
Complutense University, Madrid, Spain

1. Introduction
2. Gender Bias in Health Research and Care
3. Life Cycle: Differential Health Characteristics of Elderly Women
4. Some Cultural Differences: Interactions among Sex/Gender, Social Class, Ethnic Groups, and Culture
5. Mental and Physical Health
6. Women's Health on the Internet
Further Reading

GLOSSARY

bias Trend in the recording, analysis, interpretation, and/or revision of data leading to the conclusions drawn from them being false in some way or another.
health State of physical, social, and mental well-being.
morbidity Lack of physical or psychological well-being.
mortality Frequency of deaths in a given population during a specific time period.
prevalence Number of registered cases of a given illness divided by the number of persons in a particular population during a specific time period.
risk Any variable of personal, environmental, or biological lifestyle that might be associated with a greater likelihood of suffering a specific health impairment.

During the past 20 years or so, significant data have appeared on the differences between men and women regarding the way in which they become ill and die. This has given rise to consideration of the existence of possible gender bias in matters related to health, a bias that may have facilitated the so-called "invisibility" of women in areas of health research, professional training, and care. However, there is a differential profile of women with regard to health (both physical and mental).

1. INTRODUCTION

The worst consequences of the possible gender bias in health probably are that less effective provisions are made in community health programs, the conclusions of empirical studies are less valid, and estimates of the prevalence rates for health problems are less effective given that all of these are frequently based on what is known of illnesses and problems from the more studied sex, that is, the male one.

Gender bias in health questions is already apparent when statistical data on health are collected and published. In many countries, even developed ones, there were no differential statistics by sex on morbidity and mortality until recently, so conclusions could not be reached regarding the possible existence of differential rates or their causes.

Moreover, it is true that there is an increasingly greater awareness of this gender bias in works related to health, so there are an increasing number of studies of subjects related to gender bias and how it is exerted

©2004 Elsevier Inc.
All rights reserved.

on the medical/psychological act, on diagnoses, on treatments, on community action plans, and on how health-related matters are researched.

What all of this implies is that women are beginning to become "visible" to health professionals who increasingly recognize that there may be both differences and similarities between men and women in the ways in which they succumb to illness, in how they are cured, in how they metabolize pharmaceutical products, and in how they react to therapy in the etiology of illness.

This recognition of the differential fact has two main causes: the stubbornness of the facts (which serve as a constant indicator) and the growing feminization of health-related professions (which has led to an increase in the sensitivity of health professionals themselves to matters related to women's health).

Be that as it may, the fact is that since the late 20th century, there has been an explicit recognition that women's health is a specific area of study. Scientists, health professionals, politicians, and the like are beginning to realize that there are health-related aspects that affect men and women in differential ways (both quantitatively and qualitatively).

Because it is recognized that women do not necessarily react to preventive, diagnostic, and treatment processes for illness (both physical and mental) in a similar manner to that of men, health-related research is beginning to include, in systematic fashion, feminine subjects. Until quite recently, this was still quite infrequent.

Naturally, "women" itself is not a compact uniform unit. The health issues may vary, to a large extent, from one woman to another. This depends on variables such as a woman's stage in the life cycle, her social class, her ethnicity, her working conditions; her social and family environment, recent or distant life events, and her personality style. Many of these factors have similar effects on the health of both men and women, whereas others do not. Under the heading "women's health," one can find factors that are known or suspected to have differential effects on the health of women, to a greater or lesser extent, and the illnesses that, for one reason or another, are more related to women.

2. GENDER BIAS IN HEALTH RESEARCH AND CARE

Gender bias in health research and care can be displayed in three different ways: false assumptions of equality between the sexes, false assumptions of differences between the sexes, and reproduction-oriented limits.

2.1. False Assumptions of Equality Between the Sexes

It is assumed that problems, risks, and protecting factors in the health of men and women are the same when in fact they are not the same. A typical example is that of the efficacy of aspirin in stroke prevention. Clinical tests usually do not include a sufficient number of women to produce any conclusion (instead, an extrapolation is made from studies carried out on men), but those studies that do include sufficient women find that they do not always benefit from aspirin in preventing primary strokes. This difference between the sexes seems to be due to the existence of different stroke mechanisms in men and women.

2.2. False Assumptions of Differences between the Sexes

Biological and psychological differences also are assumed to exist when in fact what exist are similarities. Until only a few years ago, cardiovascular disease was not believed to be a "female illness." Unfortunately, many women still believe this; in a 1995 Gallup poll, only 19% of American women expected to die from cardiovascular disease, whereas two-thirds of the women interviewed expected to die from cancer. Nevertheless, in industrialized countries, cardiovascular disease is the main cause of death not just among men but also among women, and not just among elderly women but also among younger women. As a result of this gender bias, differences have been produced—and on occasion are still being produced—between men and women with respect to access to medical care. Women wait longer before seeking medical help, and even today they are less likely to receive aggressive intervention once their illness has been diagnosed. Consequently, they are more likely to die from their cardiovascular disease than are men.

2.3. Reproduction-Oriented Limits

Differential problems of women's health are assumed to be restricted to just one area: reproductive health. Although these problems are important, they are not the only ones. Thus, differential attention to women's health cannot be restricted to the areas of pregnancy, contraception, and family planning.

3. LIFE CYCLE: DIFFERENTIAL HEALTH CHARACTERISTICS OF ELDERLY WOMEN

Each stage of woman's life cycle has specific health-related aspects: pregnancy, menstruation (e.g., menstrual disorders, premenstrual syndrome), contraception, abortion, maternal morbidity, and infertility. But it may be the final stage of the life cycle that is receiving the most attention currently given the sociological changes taking place worldwide. These changes are leading to a constant increase in the number of elderly people, particularly the number of elderly women given that women have a longer life expectancy than do men (Fig. 1).

Thus, the study of health in the final stage of the human life cycle is becoming particularly important with regard to women and health. However, emphasis must be placed only on studying the fatal illnesses at this stage but also on studying the conditions and disorders that, although not fatal, have important functional consequences and affect people's quality of life. In the case of women, these conditions include matters such as menopause-related changes, functional and cognitive deterioration, osteoporosis, and urinary incontinence. The prevalence of these disorders increases considerably with advancing age, making them very important for the population of elderly women, who have a greater likelihood of surviving to an advanced age than do men.

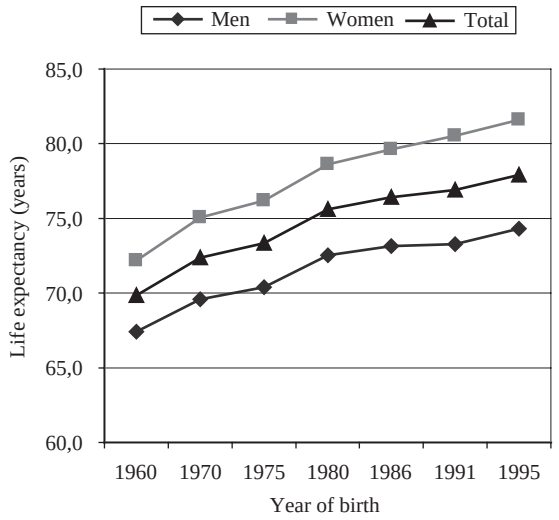


FIGURE 1 Life expectancy at birth of men and women in a developed country.

Furthermore, older people frequently suffer from chronic disorders, so it becomes essential not to assess illnesses in isolation but rather to study the global impact of these illnesses as a whole on any person. To evaluate an elderly woman's state of health, one must focus on her functional state, her incapacity, and her level of independence—and not just on a list of her illnesses—regardless of the interaction among them.

But neither should it be forgotten that recent research into healthy aging has helped to cast doubt on the stereotypes that regard aging as a period of inevitable frailty and decline. Individual differences also play an important role here. Epidemiological studies have shown that many elderly women still function at a high physical level, still maintain good cognitive abilities, and still retain a state of well-being for many years of old age. This has meant that research into the determining factors of a healthy old age is ever more frequent.

Subjects that must be tackled when dealing with the health of elderly women include menopause (which is being studied more and more), risks and benefits of hormone replacement therapy, musculoskeletal diseases such as osteoporosis (with important implications due to the risk of fractures, for its functional consequences, and for the use it requires of medical services and long-term care), arthritis (which is clearly the most frequent chronic condition in the elderly population, at least in developed countries), and urinary incontinence (with its possibly important psychological, social, and physical impacts). Sensory impairment must also be considered given that sight has a great impact on global functioning and quality of life and that impairment ranges from easily correctable cases (e.g., cataracts) to the most serious cases (e.g., macular degeneration, diabetic retinopathy, glaucoma). Also, hearing impairment heavily influences quality of life for the elderly, and its prevalence rapidly increases with age. Finally, there is loss of taste and smell, and although this has less impact on the quality of life, it is nonetheless important. For example, in relation to food, the loss of taste and smell may, on the one hand, lead to a reduction in calories ingested (because food is less appetizing) and may, on the other, hamper the ability to distinguish food that has gone "bad" (with an increasing likelihood of food poisoning). Last is senile dementia and especially Alzheimer's disease. Women themselves are not more susceptible, but because they live longer than do men, they have a greater chance of suffering from dementia and of living a long time with its symptoms.

4. SOME CULTURAL DIFFERENCES: INTERACTIONS AMONG SEX/GENDER, SOCIAL CLASS, ETHNIC GROUPS, AND CULTURE

Existing data show, logically, that the risk of suffering illnesses varies among different subgroups of women, characterized by ethnic group and/or socioeconomic level. Most studies are still descriptive, and the few that center on the explanatory level in general opt for the interaction between environmental factors (e.g., differences in the likelihood of exposure to risk factors) and biological factors (e.g., genetic differences affecting the metabolism of these risk factors),

Coherent relationships have been found between women's state of health and their ethnic and class groups. In general, the financially poorest women and those belonging to minorities have poorer health. In any case, there are exceptions to this general rule, meaning that nonsimplistic explanations must be sought for this association and attempts must be made to develop causal theories that would enable an improvement to be seen in the capacity for intervening to improve health levels.

The case of obesity may serve as an example. The frequency of this disorder has shown a steady increase during recent years, especially in countries such as the United States, where one-third of White women and half of Hispanic and African American women are considered to be obese. Moreover, there is an important inverse correlation between socioeconomic status and obesity among women, although it has not been possible to establish the causal direction of this relationship. Causality almost certainly works both ways insofar as poverty increases the risk of women becoming obese and obesity increases the risk of women's socioeconomic level declining.

Cultural differences also appear when comparisons are made between industrialized countries and developing countries. In this case, there are two typical examples: maternal mortality and violence inflicted on women.

4.1. Maternal Mortality

Although maternal mortality is a practically unknown phenomenon in the developed world, it still has tragic dimensions in developing countries. The fundamental cause of these drastic differences is related to the adequate distribution of health resources that enable factors such as uterine bleeding and other birth-related crises to be controlled.

4.2. Violence Inflicted on Women

Violence inflicted on women ranges from what is known as female circumcision (or, less euphemistically, as female mutilation) to physical and psychological ill treatment, either in the home or in areas of war and social conflict. These latter two types of violent situations are produced both in developed and poor countries, as shown by recent history and in daily newspaper headlines. But it is true that the frequency of both types may vary, depending on the social characteristics, and that the consequences may be different for both the victim and the aggressor, depending on social sensitivity to this type of aggression.

5. MENTAL AND PHYSICAL HEALTH

An example of representative data regarding predominantly physical illnesses is reported in Table I. With regard to illnesses related to the immunological system, the basis for the differential prevalence between the sexes is related to the greater response by the immunological system of men, probably due to both the differences in sexual hormones and the influence of environmental factors. Specifically, the incidence of asthma would be related, on the one hand, to the greater exposure of children to factors increasing the incidence of asthma (which would explain the lower prevalence in girls under 12 years of age) and, on the other, to the greater likelihood of women being diagnosed as asthmatics and the influence of sexual hormones on predisposition to allergy. These latter two factors would explain the greater incidence in females over 12 years of age.

With regard to mental illnesses, there is no longer any doubt that, in general, men differ from women in the prevalence and severity of mental illnesses. Table II provides a list of the major disorders, including reasons for such a differential prevalence.

6. WOMEN'S HEALTH ON THE INTERNET

Information regarding women and health on the Internet is plentiful. Three types of Web addresses on women and health can be differentiated. First are those that refer to women's studies and that consist of pages or Web sites on women and including lists (in alphabetical order) of centers, departments, and programs of

TABLE I
Differences between the Sexes in Illnesses Related to the Immunological System, Cardiovascular Illnesses, and Cancer

	Differences between sexes	*Sex–age interaction*
Illnesses related to immunological system	Hypothyroidism: between 3 and 20% for women and between 1 and 7% for men	
Thyroids	Hyperthyroidism: between 2 and 19% for women and 1% for men	
	Thyroid cancer: three times greater in women than in men	
Rheumatic arthritis	Better prognosis for women than for men	
Multiple sclerosis	Affects mostly women (65% of total population affected)	Women most affected at beginning of 20s and men at end of the 20s
Lupus	Predominance in women	From ages 15 to 45 years, the ratio between women and men is 12 to 1; in infancy and old age, the ratio is 2 to 1
Asthma	Severity greater in women; differences not too great	Before years of age, 12 lower prevalence in girls; after 12 years of age, greater incidence in females
Cardiovascular	Prime cause of death among males and females; differences in prevalence according to type of illness	Principal cause of death among younger women; accounts for half of deaths in older women
Cancer	Similar prevalences except for those related to reproduction	
AIDS	Less than half (40%) of cases found among women	

women's studies. As an extra section, there are subjects related to women's health (e.g., http://creativefolk.com and www.research.umbc.edu in the United States). In Europe, the Women's Studies EuroMap Web site (http://women-www.uia.ac.be), which offers detailed information on WISE (the European professional association for women's studies), is particularly noteworthy.

Second, are the pages of official bodies, associations, and organizations that provide information on health-related services throughout the life cycle of women. Examples include those for Canada (www.swc-cfc.gc.ca) and the United Nations (www.un.org). Also included are the pages for the Society for Women's Health Research (www.womens-health.org) and for the International Women's Health Coalition (www.iwhc.org). In Europe, one page (www.aviva.org) collates information on women in all aspects, including those that are health related. There are also examples of national pages, including those of the Women's Health Council (www.nwci.ie), which represents organizations of women throughout Ireland; the Conseil National des Femmes Luxembourgeoises (www.cnfl.lu), which collates official documents from the various international conventions on women, statements, and resolutions from bodies such as the United Nations and the European Community; and the Instituto de la Mujer in Spain (www.mtas.es/mujer/) and the Sociedad Iberoamericana de Información Científica in Latin America, (www.siicsalud.com), which provides information on health in general, and on women's health in particular, with an excellent database.

Third are the Web gateways "women and health" and the online magazines. There are both general and specific examples, and both kinds are geared to women interested in a wide variety of topics about their state of health who can find what they need there. For example, the "webofwomen" (Spanish) provides information on lifestyles, social assessment, health, key concepts in health, Web addresses, topical information, meetings of interest, and the like (www.webdelamujer.com). Some are commercial in character and include advertisements for commercial products and services for women. They also offer the opportunity to consult online with experts on various topics as well as online questionnaires on cardiovascular

TABLE II
Differences between the Sexes in Mental Disorders

	Basis of differences between sexes	*Differences between sexes*	*Sex–age interaction*
Mood disorders	Women, all things being equal, tend to seek more professional help. There are also differences in the response to treatment	Higher rates in women	Differences increase during puberty
Anxiety disorders	Women show higher levels of arousal, greater psycho-physiological response to stress, higher frequency of somatic symptoms, and greater sensitivity to anxiety symptoms	High rates in women; one-third have some type of attack throughout their lives	
Posttraumatic stress	Differences between sexes appear when this syndrome is related to violence and a particular moment in the life cycle.	Frequency two times higher in women than in men; duration of symptoms four times higher in women than in men	At the end of adolescence and the beginning of adult life, greater risk of suffering this type of violence
Eating disorders	Biological vulnerability exists in integrating and tackling the following factors: • Sociological factors: role of woman in cultural expectations regarding weight and physical appearance • Psychological factors: characteristics of personality, family functioning, and stressful life events	• Anorexia: 0.28% average incidence among young women • Bulimia: 1% average incidence among young women	Increase of incidence among young women between 15 and 24 years of age and from 50 years of age onward
Addictive disorders		Higher in men than women.	

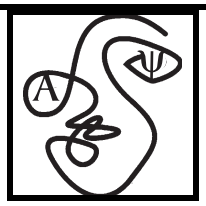
symptoms, depression, anxiety, and the like. Among them are pages devoted exclusively to certain of these topics, for example, one devoted to pregnancy and its circumstances (http://womenshealth.about.com).

See Also the Following Articles

Aggression and Culture ■ Gender and Culture ■ Gender Role Development ■ Prejudice and Discrimination

Further Reading

Goldman, M. B., & Hatch, M. C. (2000). *Women and health*. San Diego: Academic Press.
Pregler, J. P., & DeCherney, A. H. (2002). *Women's health principles and clinical practice*. San Diego: Academic Press.
Shimp, L., & Smith, M. (2000). *Twenty common problems in women's health care*. New York: McGraw–Hill.
Veet, L. L. (2000). *Women's health care case studies*. San Diego: Academic Press.
Zerbe, K. J. (1999). *Women's mental health in primary care*. San Diego: Academic Press.



Work Adjustment

Beryl Hesketh
University of Sydney, Sydney, New South Wales, Australia

1. Structural Aspects of Work Adjustment
2. Work Adjustment Process
3. Work Adjustment Assessment
4. Work Adjustment Counseling
Further Reading

GLOSSARY

adjustment style Modes of adjustment that individuals use to achieve a better fit with their work.

correspondence The match between the skills, abilities, interests, and needs of the individual and the skill and ability requirements of the job as well as the rewards offered by it.

person–environment fit A summary statement of the match between a person and the environment, typically used in reference to a work environment.

Minnesota Theory of Work Adjustment (TWA) A comprehensive structural and dynamic framework used to explain how individuals maintain and improve their job satisfaction and satisfactoriness.

work adjustment Ongoing process by which employees and employers seek to achieve and maintain a fit between their respective needs and supplies.

Work adjustment is a term that is used to describe a desired state of correspondence or fit between an individual's capabilities and wants and the commensurate requirements and rewards or opportunities provided within an organization. The term can also refer to the processes used by individuals and their employers to achieve work adjustment. Work adjustment has been used widely within the rehabilitation and vocational literature. The Minnesota Theory of Work Adjustment originated from research undertaken for the Rehabilitation Services Administration but has gained status as a major general fit theory in vocational and industrial/organizational psychology. In the rehabilitation context, work adjustment is part of the recovery process following illness, accidents, or other trauma that severely affect performance capabilities and limitations, requiring an adjustment to a different set of work expectations. In a more general context, work adjustment becomes important whenever change occurs within the individual or in the organizational requirements. Technological change, globalization, mergers, takeovers, and other external events have increased the requirement for ongoing work adjustment and adaptive performance. Change may also occur as a part of life span development, such as in transitions from school to tertiary study and work, transitions from work to unemployment, and retirement transition. Although this article focuses on a narrower meaning of work adjustment, the term in its broader sense incorporates a range of concepts, such as adaptive performance, coping behavior, change self-efficacy, role-breadth self-efficacy, ergonomic interventions, selection, and training.

©2004 Elsevier Inc.
All rights reserved.

1. STRUCTURAL ASPECTS OF WORK ADJUSTMENT

The Minnesota Theory of Work Adjustment (TWA) provides a useful framework for examining work adjustment and relating the concept to other theories and components of industrial/organizational and vocational psychology. In its original form, TWA outlines a structural component that describes people and work environments in terms of the correspondence between the knowledge, skills, and abilities (KSAs) required in a job and those provided by the individual, giving rise to satisfactory performance. This is the performance component of work adjustment, which is valued from an employer perspective. However, it also contains a motivational component that is dependent on the correspondence between the needs, values, and interests of the individual and the corresponding rewards and reinforcers provided by the job, giving rise to satisfaction. Continued interaction (tenure) between an individual and the job/organization is a function of how satisfied an individual is and how satisfied the employer is with the individual's performance. Higher levels of satisfaction and performance contribute to an increased period of interaction between the parties (tenure). Dissatisfied individuals do not necessarily perform less well, but they are more likely to leave a job. In most organizations, if an individual is not performing as required, the appointment is terminated.

There are numerous studies directly testing the main propositions of fit in TWA, but studies on employee selection also represent a large body of research relevant to the relationship between the required KSAs and those supplied by the individual, a major component of TWA. Validity studies of ability or personality dimensions that take into account "fit" with job requirements when choosing appropriate predictor and criterion dimensions are more compatible with TWA than are those that focus only on the relationship between predictors and performance without the moderating influence of job requirements.

In terms of the motivational component of the TWA, much vocational research has emphasized the match between individual needs, interests, and values and outlets for these as a predictor of job satisfaction and eventually turnover. In addition to the work value and ability (KSA) dimensions used in the TWA, other constructs can be used as a basis for commensurate assessment of people and their environments in relation to both the performance component and the satisfaction and motivation component. Holland's hexagonal theory of interests (realistic, investigative, artistic, social, enterprising, and conventional) provides one such example, as do multiattribute test batteries and the Five Factor model of personality. Tracey has outlined a spherical model of interests, the Personal Globe Inventory, and Gustaffson and Mumford and also De Fruyt argue for fit on the basis of a configurational or person-centered approach.

2. WORK ADJUSTMENT PROCESS

Adjustment at an individual level can be likened to adaptation at the species level. Dawis and Lofquist outline style variables that describe the extent to which discorrespondence is tolerated by both individuals and organizations (flexibility), the type of action initiated to improve correspondence (active or reactive modes of adjustment), and the length of time during which the parties continue to make an effort to achieve adjustment before terminating the relationship (perseverance). An active mode of adjustment relates to the activities undertaken by the individual to change the job (e.g., negotiating to eliminate aspects of the work that are too difficult or not enjoyable) or to activities the employer initiates to change the individual (e.g., training courses to provide the employee with the desired skills). Reactive modes of adjustment relate to changes made to the self. For example, in the case of individuals, they may recognize that their expectations are unrealistic and hence modify their work-related goals and expectations. From the perspective of the environment, an employer may redesign the job to take better account of the performance capabilities and limitations of the individual. Active and reactive processes of adjustment are continuous, and together aim to achieve better work adjustment.

In the recent literature in industrial and organizational psychology, terms used that are similar to the style variables in the TWA include flexibility, adaptability, and coping behavior. Perseverance in achieving adjustment is related to the extent to which an individual believes that he or she has the skills and capacities to make the necessary active and reactive adjustments.

3. WORK ADJUSTMENT ASSESSMENT

Determining the extent to which an individual is adjusted to a work situation involves discerning the fit between the individual's performance capabilities and those required

by the job as well as the fit between the individual's needs and the rewards of the work. Measures used draw on the individual differences literature for assessing cognitive and other skills, personality, interests, needs, and values. Job analysis techniques are also needed to analyze jobs to determine job requirements. This level of assessment provides a diagnostic basis for helping individuals achieve a better fit. Outcome measures used to assess work adjustment include job satisfaction, performance, and tenure. Satisfaction can be considered as part of the broader family of constructs related to mental health and well-being, whereas recent models of performance highlight the need to consider task, contextual, and adaptive performance. Tenure, the ultimate assessment of work adjustment, is best considered in the broader context of withdrawal behavior—a concept that includes absenteeism, tardiness, and early retirement as well as turnover.

4. WORK ADJUSTMENT COUNSELING

The processes used to help individuals achieve better work adjustment derive from the core constructs discussed previously. These constructs are discussed explicitly by Lofquist and Dawis, who address additional areas such as relationship counseling in which the concepts of correspondence and adjustment are also important. In work adjustment counseling, placement in a job that offers a good fit from both the individual and the employer perspective is the primary aim. However, it is seldom possible to achieve adjustment and fit purely on the basis of selection and placement. Counselors can work with individuals and their employers to try to encourage realistic expectations of each other and to help ensure that the required training and ergonomic or job redesign interventions are undertaken to facilitate better fit. All career counseling includes elements of work adjustment counseling, but the term tends to be reserved for use in the rehabilitation of employees following injury or other trauma.

See Also the Following Articles

Career Counseling ■ Coping ■ Holland's Theory (Vocational Personality Types) ■ Person–Environment Fit ■ Structure of Interests ■ Vocational Interests

Further Reading

Allworth, E. A., & Hesketh, B. (1999). Construct-oriented biodata: Capturing change related and contextually relevant future performance. *International Journal of Selection and Assessment*, *7*, 97–111.

Bretz, R. D., & Judge, T. A. (1994). Person–organization fit and the theory of work adjustment: Implications for satisfaction, tenure, and career success. *Journal of Vocational Behavior*, *44*, 32–54.

Dawis, R. V. (1994). The theory of work adjustment as convergent theory. In M. L. Savickas, & R. W. Lent (Eds.), *Convergence in career development theories: Implications for science and practice*. Palo Alto, CA: CPP Books.

Dawis, R. V., & Lofquist, L. H. (1984). *A psychological theory of work adjustment*. Minneapolis, MN: University of Minnesota Press.

De Fruyt, F. (2002). A person-centred approach to P–E fit questions: Using multiple-trait model. *Journal of Vocational Behavior*, *60*, 73–90.

George, J. M., & Jones, G. R. (1996). The experience of work and turnover intentions: Interactive effects of value attainment, job satisfaction, and positive mood. *Journal of Applied Psychology*, *81*, 318–325.

Gustafson, S. B., & Mumford, M. D. (1995). Personal style and person–environment fit: A pattern approach. *Journal of Vocational Behavior*, *46*, 163–188.

Hanisch, K. A., & Hulin, C. L. (1991). General attitudes and organizational withdrawal: An evaluation of a causal model. *Journal of Vocational Behavior*, *39*, 110–128.

Hesketh, B., & Gardiner, D. (1993). Person–environment fit: A reconceptualization and empirical test. *Journal of Vocational Behavior*, *42*, 315–332.

Holland, J. H. (1997). *Making vocational choices: A theory of vocational personalities and work environments*. Englewood Cliffs, NJ: Prentice Hall.

Lofquist, L. H., & Dawis, R. V. (1991). *Essentials of person–environment correspondence counselling*. Minneapolis, MN: University of Minnesota Press.

Schmitt, N., & Chan, D (1998). *Personnel selection: A theoretical approach*. Thousand Oaks, CA: Sage.

Tracey, T. J. G. (2002). Personal Globe Inventory: Measurement of the Spherical Model of Interests and Competence Beliefs. *Journal of Vocational Behavior*, *60*, 113–172.

Warr, P. (1987). *Work, unemployment, and work adjustment counseling mental health*. Oxford, UK: Oxford University Press.

Work and Family, Relationship between

Jeffrey H. Greenhaus
Drexel University, Philadelphia, Pennsylvania, USA

Romila Singh
University of Wisconsin, Milwaukee, Wisconsin, USA

1. Introduction
2. Definitions of Work and Family
3. Mechanisms Linking Work and Family
4. The Effect of Work on Family Life
5. The Effect of Family on Work Life
6. Gender Differences in Work–Family Relationships
7. Organizational Actions to Help Individuals Manage Work–Family Relationships
8. Individual Actions to Manage Work–Family Relationships
9. Conclusion
Further Reading

GLOSSARY

family An emotional unit based on love and affection that provides psychological security and nurturance to its members.

work An instrumental activity that is intended to provide goods and services to support life or produce something of value for others.

work–family accommodation The process by which individuals reduce their involvement in one role to accommodate the demands of the other role. It also refers to a strategy that can be used in response to actual or anticipated work–family conflict.

work–family balance The extent to which individuals are equally involved in—and equally satisfied with—their work role and their family role.

work–family compensation Efforts to offset dissatisfaction in one role by seeking satisfaction in another role. Compensation may be either reactive or supplemental.

work–family conflict Simultaneous pressures from both work and family roles that are mutually incompatible in some respect such that meeting the demands of one role makes it difficult to meet the demands of the other role.

work–family enrichment The process by which one role strengthens or enriches the quality of the other role.

work–family segmentation The intentional separation of work and family roles such that the thoughts, feelings, and behaviors of one role are actively suppressed from affecting the individual's performance in the other role.

work–family spillover The transfer or application of skills, values, emotions, and behavior from one role to the other role. Spillover may be either positive (helpful) or negative (harmful).

Work and family lives are interconnected in many ways. Work experiences can influence family life, and family experiences can affect work life. Some of these effects are negative (work–family conflict), and other effects are positive (work–family enrichment). In this article, different mechanisms by which work and family lives influence each other, gender differences in work–family relationships, and organizational and individual actions to reduce work–family conflict and to promote work–family enrichment are discussed.

©2004 Elsevier Inc.
All rights reserved.

1. INTRODUCTION

Scholars and journalists write extensively about the difficulties of juggling work and family commitments in contemporary society. Individuals clamor for information about how to cope with extensive work and family responsibilities, and employers increasingly offer work–family programs to help their employees relieve work–family stress and, at the same time, meet their obligations in the workplace. Previously thought to be separate worlds that do not impinge on one another, work and family lives are currently recognized as being inextricably intertwined, sometimes for the better and sometimes for the worse.

What lies behind this interest in work–family relationships? First, the increasing representation of women in the workforce—especially mothers of young children—has rendered the dual-earner family a dominant family structure in the United States. With the blurring of traditionally rigid gender roles—women as full-time homemakers and men as sole providers—both partners in a dual-earner relationship are confronted with the daily challenge of handling their work and family responsibilities in a way that meets the needs of the family and the employer alike.

Moreover, a divorce rate that has stabilized at approximately 50% in the United States has produced increasing numbers of single parents in the workforce. These single parents, most of them women, often struggle to meet extensive work and family responsibilities, often without a great deal of support from their families or their employers.

In addition to, or perhaps because of, these structural changes in family life and workforce participation, employees' values have changed in recent years. Women and men are seeking balance in their lives in deciding what jobs to accept and what jobs to leave. Job applicants are increasingly raising the issue of work–family balance in their employment interviews, and a sizeable number of employees seem willing to make sacrifices in their careers to achieve a high quality of life. Some individuals have chosen to leave their profession when it continually interferes with their family lives.

All these forces have combined to produce an intense interest in the relationship between work and family lives. Researchers in psychology, sociology, and management have risen to the challenge. The past three decades have witnessed an explosion of research on the intersection of work and family roles. Understanding how work can affect family—and how family can affect work—has been the aim of most of this research. Gaining insight into the relationships between work and family lives is important for a number of reasons.

First, the more that is learned about the intricate relationships between work and family commitments, the more likely employers can provide relevant programs to help their workers manage their work and family lives. Second, understanding the interplay between work and family roles can help women and men learn how to obtain employment that is compatible with their desired lifestyle, how to cope with the stresses of daily life, and how to gain the support of family, friends, and coworkers in relieving work–family stress.

In this article, a wide range of issues regarding the ways in which work and family lives affect one another are discussed. First, the two key roles—work and family—are defined. Then, a variety of mechanisms that have been proposed to explain the relationship between work and family roles are discussed. More detailed discussions of the effects of work on family life and the effects of family on work life are then presented. Moreover, because gender is so central to the interplay between work and family roles, gender issues in work–family relationships are discussed. Finally, some steps that employers and individuals can take to manage the work–family interface effectively are presented.

2. DEFINITIONS OF WORK AND FAMILY

In order to explain the interdependence between work and family roles, it is first important to define the terms work and family. Work has most commonly been defined as an instrumental activity that is intended to provide goods and services to support life or produce something of value for others. Most of the research in this area has focused on paid work or employment.

Researchers have defined family most often in terms of the presence of a spouse, children, or kin in the household. A more inclusive definition expands the notion of family beyond marital, adoptive, or biological ties to include people related by affection, obligation, dependence, or cooperation. Thus, a family is considered an emotional unit based on love and affection and one that provides psychological security and nurturance to its members. Moreover, the United States has witnessed an increasing multiplicity of family structures, such as two-parent families, one-parent families, cohabiting couples, gay and lesbian families, and

extended family households. Nevertheless, much of the research on the work–family interface has been conducted on samples of two-parent families.

3. MECHANISMS LINKING WORK AND FAMILY

Researchers have identified several mechanisms that help explain the relationship between work and family roles. Early research focused on the ways in which work and family lives conflicted or interfered with one another. Over the years, other explanations emerged that explained a more diverse array of linkages between work and family lives. The most prominent mechanisms that are believed to explain the dynamics of the work–family interface are conflict or interference, accommodation, enrichment, spillover, balance, compensation, and segmentation.

An examination of the work–family literature reveals the consistent popularity of the work–family conflict perspective as an explanation of the manner in which work and family roles influence one another. Work–family conflict refers to simultaneous pressures from both work and family roles that are mutually incompatible in some respect such that meeting the demands of one role makes it difficult to meet the demands of the other role. Work–family conflict has been further distinguished into three specific types of conflict: time-based conflict, strain-based conflict, and behavior-based conflict. Time-based conflict occurs when devoting substantial time to one role makes it difficult to meet the demands associated with the other role. Strain-based conflict occurs when stress arising in one role is carried into the other role and affects individuals' ability to meet the demands of that role. Behavior-based conflict occurs when behavior that is developed in one role is incompatible with the expectations regarding behavior in another role.

Because work–family conflict represents interference between the two roles, researchers have also examined the direction of this interference. Specifically, work-to-family conflict occurs when work demands interfere with the quality of family life. On the other hand, family-to-work conflict occurs when family demands and pressures interfere with the fulfillment of work responsibilities. The consequences of work–family conflict on individuals' work and family lives are examined later.

Another mechanism that has been used to explain the relationship between work and family roles is that of work–family accommodation. Work–family accommodation is a strategy that can be used in response to actual or anticipated work–family conflict. Accommodation refers to the process by which individuals reduce their involvement in one role to accommodate the demands of the other role. Researchers believe that the reduction in involvement usually occurs in the role that is less important to the individuals and can take two forms. Reduction in behavioral involvement in a role occurs when individuals curtail the amount of time devoted to a role, such as cutting back on the time devoted to work in order to meet their family responsibilities more effectively. Reduction in psychological involvement occurs when individuals restrict their level of ego attachment to a particular role or reduce the importance attached to that role, such as when individuals display less involvement in their family lives in order to fulfill their job responsibilities more effectively.

Whereas work–family conflict represents the negative effects of one role on the other role, work–family enrichment refers to the process by which one role strengthens or enriches the quality of the other role. Work and family roles each provide a variety of resources—tangible and intangible—that can enhance the quality of the other role. Subsequent sections discuss how specific resources from work and family roles have the capacity to provide positive experiences in the other role.

Work–family spillover most often refers to the transfer or application of skills, values, emotions, and behavior from one role to the other role. Work–family spillover can originate in either of the two roles (i.e., spillover can be either from work to family or from family to work). Furthermore, researchers have distinguished two forms of spillover—positive spillover and negative spillover. With positive spillover, the transfer or application of skills, values, emotions, and behavior from one role promotes an individual's performance or satisfaction in the other role, thereby producing work–family enrichment. Negative spillover reduces individuals' performance or satisfaction in the other role, producing work–family conflict. For example, individuals' satisfaction at work is believed to enhance their family life (positive spillover) and, conversely, job stress is likely to create stress in their family interactions (negative spillover).

Presumably, individuals who lead a balanced life can reduce work–family conflict and increase work–family enrichment. Although there is no shortage of advice on how to achieve work–family balance, relatively little research has been conducted on balance or its consequences. We view work–family balance as the extent to which individuals are equally involved in—and equally satisfied with—their work roles and family roles. It has

been suggested that individuals who are balanced in their orientation gain a great deal from work and family activities, experience less stress in carrying out their roles, and derive high self-esteem from the competence they achieve in their work and family pursuits.

Work–family compensation represents efforts to offset dissatisfaction in one role by seeking satisfaction in another role and as such suggests that work and family experiences tend to be antithetical. Researchers have distinguished two forms of compensation. First, individuals may choose to decrease involvement in a dissatisfying role and increase involvement in a more satisfying role. Alternately, individuals may respond to dissatisfaction in one role by pursuing rewarding or fulfilling experiences in the other role. The latter form of compensation has been further distinguished into supplemental and reactive compensation. Supplemental compensation occurs when individuals shift their pursuits for rewarding experiences from the dissatisfying role to a potentially more satisfying one. For example, individuals with little autonomy at work may seek autonomy outside of the work role. Reactive compensation occurs when individuals seek to redress negative experiences in one role by pursuing contrasting experiences in the other role. Pursuing leisure activities after work or resting at home due to fatiguing work are examples of reactive compensation.

Although the mechanisms discussed in this section describe how work and family can affect one another, there is another perspective that focuses on the independence of the two roles. Work–family segmentation originally referred to the notion that work and family spheres exist side by side such that individuals can participate in one sphere without any influence on the other sphere. This notion of separateness (or independence) between work and family lives in time, space, and function implies the compartmentalization of one's life into distinct and separate worlds. Recently, work–family segmentation has referred to the intentional separation of work and family roles such that the thoughts, feelings, and behaviors of one role are actively suppressed from affecting the individual's performance in the other role. Segmentation is now believed to be a process of coping with the stress from either the work role or the family role. As such, it is viewed as a deliberate strategy to inhibit the occurrence of interference between domains and also as a way to maintain a boundary between one's work and family roles.

The previous brief description of mechanisms that explain the nature of the relationship between work and family roles provides the foundation for discussing in more detail the different ways in which work and family lives affect one another.

4. THE EFFECT OF WORK ON FAMILY LIFE

Work experiences can affect family life in many different ways. Some of these effects are negative, in which work may be considered an "enemy" of the family. In other respects, however, work can be an "ally" of the family, enriching the quality of family experiences.

4.1. Negative Effects of Work on Family Life

Much of the early research on the work–family interface was based on a scarcity hypothesis that assumes that time and energy are fixed commodities, and that people who participate in a multitude of social roles (such as work and family) inevitably experience conflict between the roles. Many individuals' job responsibilities are so time-consuming or so inflexible that they interfere with a full and satisfying family life. Moreover, as a result of this time-based conflict, individuals can become dissatisfied with their marriage and their family and may experience substantial depression and dissatisfaction with their lives.

Although the shortage of time to devote to both work and family activities is part of the problem, another issue, perhaps as significant as time, is the psychological involvement that some employees invest in their work and career. There is a segment of the population for whom work is the central part of their lives. These individuals identify so strongly with their work and are so absorbed in their work that they have little left to give to their family. They may be so preoccupied with work that even when they are physically present at home, their thoughts lie at their workplace. An imbalanced, overinvolvement in work, sometimes referred to as workaholism, may have negative effects on individuals' relationships with their family and can detract from the quality of these individuals' lives.

Beyond the time devoted to work and the psychological absorption in work that can interfere with family, job stress can have negative consequences for employees' family life. Work-related stress can have many underlying causes—a boring job, too much responsibility, interpersonal conflict with a colleague, undue pressure for productivity, or the threat of job loss. Regardless of its

source, the tension, irritability, fatigue, or anxiety produced by job stressors may be "taken home" and interfere with one's performance or enjoyment within the family domain. This negative emotional spillover can render individuals in stressful jobs vulnerable to problems at home. Such individuals have trouble "turning off" the stress when they cross the boundary from work to family.

One important by-product of job stress that plays out in the family domain is the crossover of stress that may occur from one spouse or partner to the other. When individuals' job-related stress spills over into the home, the stress may be transmitted to their partners. For example, individuals who feel burned out at work may become anxious, tense, and irritable at home. These individuals' partners may begin to perceive the individuals' tension and, out of empathy toward the partners, become stressed themselves. Burned-out employees may, out of frustration, behave harshly toward their partners, which triggers stressful reactions in the partners. It is reasonable to expect that the increased stress experienced by the partner will arouse even more stress in the already anxious employee, thereby creating a vicious cycle.

Stressful work may also have negative effects on children. Individuals who bring home extensive stress from their work may withdraw from interactions with their children, act coldly toward them, and experience conflict dealing with their adolescents. One source of parental stress—job insecurity—has been found to be associated with children's negative work attitudes, low self-confidence, cognitive difficulties, and poor grades in school.

In addition to the family problems produced by time, overinvolvement, or stress within the work role, there is another way in which work lives can interfere with family. Behavior that is learned or reinforced at work may be inappropriately applied in the family domain, thereby producing behavior-based conflict. For example, individuals may believe that the pursuit of career success requires them to be detached and distant from subordinates so they can objectively appraise their subordinates' performance and make difficult decisions about their future careers. They may also believe that being highly directive toward their subordinates—telling them exactly how to perform their jobs—is an effective leadership style. However, individuals who apply the same detached, distant, and directive style at home may find that their spouse and children are understandably displeased. Therefore, the inability to "switch gears" when crossing the border from work to home can produce tension and anger in family members that detract from the quality of family life.

In summary, the negative effects of work on family life reflect one form or another of work-to-family conflict due to extensive and inflexible work hours, overinvolvement in work, job stress that carries over into the home environment, or work-related behaviors that are inappropriately applied within the family domain. Such conflict can produce distress within the family domain, withdrawal from family responsibilities, and negative parenting, and it can adversely affect one's quality of life.

4.2. Positive Effects of Work on Family Life

Although work experiences can interfere with home and family responsibilities in a variety of ways, they can also enrich the quality of family life. Therefore, in the right circumstances, work can be an ally of the family. Employment potentially provides a number of resources that can be applied fruitfully to the family domain.

One of the most powerful resources provided by a job is income. A high income can enhance a family's standard of living, provide a stimulating and healthful home environment for children, and enable parents to spend more time with children through the purchase of outside help to perform some household chores. As observed frequently in everyday life, money does not guarantee happiness, but it can serve as a resource to enhance the quality of a family's life.

However, resources derived from work go beyond money to include more intangible assets that are associated with employment. For example, individuals who are satisfied with their job tend also to be satisfied with their family life. Just as extensive job stress can spill over to the family in a negative way, good feelings about work can be transferred to the home environment through positive spillover, so an individual can be more relaxed and participate more fully in family life.

One especially important resource at work is the amount of autonomy or freedom provided on the job. Employees who experience substantial job autonomy provide a stimulating home environment for their children, tend to have more positive interactions with their children, and report that their children have relatively few health and behavior problems. Moreover, individuals with considerable job autonomy experience low levels of work–family conflict, probably because they have control over the timing and location of their work. Similarly, flexible work schedules enable individuals to gain more control over their work responsibilities, reducing their level of work–family conflict and increasing their satisfaction with family life.

Another important resource that can be garnered at work is the acquisition of skills, information, and behavior that can be applied effectively at home—the opposite of behavior-based conflict. For example, employees may learn that listening attentively to coworkers or clients is an effective work style that, when used at home, improves the quality of their relationship with other family members. Managers may realize that letting subordinates learn from their own mistakes is a technique that can be used at home to promote their children's self-esteem and independence. Moreover, employees may find that their self-esteem is bolstered when they perform well on the job. This heightened sense of self-esteem can serve as a resource because it may enable employees to interact more confidently with others in the family domain.

Finally, the social support provided by others in the workplace is a significant resource to many individuals. Employees who receive substantial support from their coworkers or supervisors tend to experience relatively little work–family conflict and enjoy greater well-being in the family domain. Sometimes, this support is based on informal relationships and friendships with other people who can provide understanding, information, and advice on juggling work and family responsibilities. In other cases, support is provided through an employer's institutionalized "family-responsive" practices and a culture that respects an employee's life outside work. Individuals who work for supportive organizations have been found to experience relatively little conflict between their work and family responsibilities and are more satisfied with their personal growth and development.

5. THE EFFECT OF FAMILY ON WORK LIFE

The previous description illustrated the numerous ways in which work experiences affect family life. Because work and family roles are mutually interdependent, family life and experiences can also affect an individual's workplace behavior, attitudes, performance, and career decisions. Just as work can either enhance or detract from the family, there are parallel positive and negative influences of family experiences on work life.

5.1. Negative Effects of Family on Work Life

Much of the early research on the effects of family on work was based on gendered work–family models that examined the influence of women's family roles and responsibilities on their work choices and performance. With the growing recognition of men's involvement in family life, scholars have begun to offer a more balanced (or gender-neutral) explanation of the effects of family on work.

A large body of the work–family literature has focused on the effects of time spent in family activities on a variety of work attitudes and decisions. The underlying premise is that both work and family roles make myriad demands on individuals' time, and in trying to meet those demands conflict inevitably results. Certain facets of one's family life are likely to make more demands on individuals' time than others. One dimension of family life—marriage—may impose only limited time demands and hence exercise a limited influence on work behaviors and performance. However, an increase in time demands may be introduced in the form of spousal job commitments and attitudes that influence a variety of career decisions, such as work hours, involvement, and withdrawal. For example, individuals whose spouses are unavailable or reluctant to share in the household chores or other family responsibilities may end up shouldering more of these responsibilities with the result that they have little time or energy left to give to work.

One particular facet of family life that has received a great deal of attention in terms of its effect on work is parental responsibilities. Parenting undeniably introduces new time and energy demands on individuals and has been shown to affect certain employment decisions, such as seeking part-time work, self-employment, and employment in family supportive organizations. Once employed, parents with young children are likely to spend relatively fewer hours on the job than nonparents and are likely to make adjustments to their work schedules that may often result in reduced time at work. Furthermore, parents may often turn down career development opportunities that involve travel in order to spend more time with their families.

Another characteristic of caregiving that taps into employees' already stretched time and energy resources is care for the elderly. With the progressive aging of the population, it has been estimated that approximately 20–30% of all workers are likely to be involved in providing some form of assistance to an older person. Employees who have extensive elder care responsibilities may limit their behavioral involvement at work by cutting back on their work hours, rearranging their work schedules, turning down relocation opportunities, and taking unpaid leave. When these responsibilities

become very demanding, it is not uncommon for some employees to withdraw completely from the workforce.

Different caregiving responsibilities demand not just physical and behavioral involvement but also require a great deal of psychological involvement in the family role with the result that individuals have little left to give to their work role. Some employees who are behaviorally and psychologically focused on their families have considerably lower aspirations for senior management positions than those who are either career focused or both career and family focused. Other individuals who are deeply involved with their family lives may choose to scale back their involvement in work, which may eventually affect their potential for getting jobs with greater autonomy or other career-enhancing assignments.

In all these instances, individuals seek to accommodate the increase in caregiving responsibilities and their associated time and energy demands by limiting their participation and investment in the work role. Although some individuals voluntarily scale back their involvement in the work role, it is also possible that some organizations are less willing to invest in employees who are extensively involved in their family lives. A full discussion of organizational responsiveness to work–family issues is presented later.

Extensive obligations to fulfill parental and caregiving roles may make individuals irritable, anxious, or tense and create disruptive family dynamics that lead to family distress. Such negative emotions and distress generated within the family role may spill over into the work domain and reduce individuals' work productivity and work attitudes.

In summary, different facets of individuals' family life, such as spousal behavior and attitudes, extensive caregiving responsibilities, and intensive involvement with family activities, can conflict with their work lives. Such conflict can limit individuals' career choices and involvement and negatively affect their work involvement, job satisfaction, aspirations for senior management positions, and their intentions to continue employment.

5.2. Positive Effects of Family on Work Life

Home and family life can also strengthen, expand, or enrich the quality of one's work life, and in this sense family becomes an ally of work. Individuals living in a family can draw on a variety of supportive resources that directly or indirectly affect their work lives.

Family members—one's spouse, children, and other relatives—are perhaps the major source of support that can either prevent family stress or buffer the individuals from the pressures of home life and enable them to participate more actively and effectively in work life. This support can take various forms, such as providing direct physical help with various household tasks or child care activities and/or providing emotional support in the form of affection, advice, affirmation, and information that helps individuals cope more effectively with various work-related problems. By acting as a sounding board and offering suggestions and feedback to their spouses, a supportive environment can be created that enables individuals to devote more time to the pursuit of important career goals.

Supportive family members, especially one's spouse, who participate more actively in family life and provide emotional support allow the individuals to spend more time and energy at work, make fewer adjustments to their work schedules, avail themselves of career-enhancing opportunities, and experience greater satisfaction with their careers than those who receive little support. Support from one's spouse not only enables individuals to cope more effectively with their work lives but also can reduce the extent to which their family interferes with work. This, in turn, can free up individuals mentally, emotionally, and physically to channel greater effort and energy into their work.

Working spouses can also provide support in the form of financial resources that can augment their personal efforts and thus create a more satisfying family environment. In theory, the security provided by spousal income can enable individuals to leave a dissatisfying job and pursue a career that is more in line with their values and aspirations.

It is not just spouses who provide significant sources of support to individuals. Children, although demanding on working parents, can also be a source of help to them. It is not uncommon for children, particularly older ones, to help out with household work and look after their younger siblings. For parents who are already stretched for time and energy, such support is valuable in enabling them to increase their time commitment to work.

Supportive family relationships can foster emotional closeness, bring peace of mind, and produce general satisfaction with one's family life. Positive feelings and fulfillment generated in one's home environment can spill over into one's work life, producing satisfying feelings toward one's work and allowing individuals to apply themselves fully at work. Positive spillover can also occur when individuals transfer the skills, knowledge, and perspectives that they acquired from their home experiences to their work lives. For

example, parenting skills such as active listening, expressing concern, and providing feedback may enhance an individual's effectiveness at work.

In summary, a supportive family environment may offset some of the negative effects of family on work and also strengthen, enhance, and enrich an individual's work life in a variety of ways. However, to fully appreciate the intricacies of work–family relationships, one must take into consideration whether these influences are similar for both men and women.

6. GENDER DIFFERENCES IN WORK–FAMILY RELATIONSHIPS

Despite a weakening gender-based segregation of roles, men and women continue to spend their time differently, especially when there are children in the family. Fathers generally devote more time to paid employment than do mothers, and mothers spend more time on household chores and caring for the children than do fathers. Moreover, women generally feel more responsible for the physical and emotional well-being of their children than do men. Given these persistent gender differences in work and family role orientations, it is reasonable to ask whether the interdependence of work and family roles is different for women and men.

Considering first the negative relationships between work and family roles, some studies have found that women experience more work–family conflict than do men. The results of these studies are consistent with the reality that employed women typically spend more total time on work, home, and family activities than do men. The number of hours that women work is rarely accompanied by a commensurate reduction in the time they devote to home and family. For this reason, it has been suggested that women work a "second shift" when they come home after a day's employment. Therefore, when women experience more work-to-family conflict than men, it is likely because they have more family responsibilities with which work can interfere.

There is also a ready explanation for those situations in which women experience more family-to-work conflict than men. Women with children, especially young children, are more likely to reduce their involvement in work in order to accommodate their family's needs than are men. This accommodation may take the form of cutting back on work hours, restricting the amount of weekend or evening work, or curtailing extensive travel. Any of these forms of accommodation may cause a woman to perceive that her family responsibilities have limited or interfered with her work-related pursuits.

However, many other studies have found that women and men experience a similar level of work–family conflict. Again, there are several explanations for this finding. First, it is possible that gender differences in work and family responsibilities have eroded to the point where women and men basically face the same degree of conflict in juggling their life roles. More plausibly, however, women may be more likely than men to choose jobs that are less demanding and therefore less likely to produce extensive work–family conflict. Historically, women have been more likely than men to base their selection of a particular occupation on its compatibility with extensive family responsibilities. Moreover, as noted previously, many women with young children have already reduced their work involvement to alleviate the daily conflicts between family and work. It is also possible that women are more effective than men at coping with work and family pressures, thereby keeping work–family conflict at a manageable level. All of these factors may explain why many studies have observed no gender difference in work–family conflict.

There has not been a great deal of research on whether positive relationships between work and family roles are different for men and women. It seems that both women and men can benefit from the support provided by their spouses, supervisors, or coworkers. However, some research suggests that the positive effect of work experiences on family life may be somewhat different for men and women. For example, women may be more likely than men to use their work experiences (especially autonomy and social relationships) to benefit their families. In addition, women respond more positively than men to the presence of a flexible work schedule and place more importance than men on child care support and alternative work hours. Despite these differences, it is clear that women and men are capable of using resources from a particular role to improve life in the other role.

7. ORGANIZATIONAL ACTIONS TO HELP INDIVIDUALS MANAGE WORK–FAMILY RELATIONSHIPS

As indicated previously, employees can experience conflict between their work and family lives. Although women continue to face extensive pressures in juggling their home and work demands, it is evident that men are increasingly involved in these same struggles. As these

challenges intensify in the future, employees will continue to demand opportunities that help them attain a balance between their work and family roles and lead more fulfilling lives. In response to changing demographic trends and employee preferences, organizations are offering a variety of programs that help employees address their work–family concerns. These programs can be broadly classified to include practices that focus on dependent care issues and others that focus on flexible work arrangements.

7.1. Dependent Care Arrangements

As previously discussed, a large number of working parents often scale back their time at work, make significant adjustments to their work schedules, miss work, and curtail their involvement at work in order to fulfill some of their child care responsibilities. Organizations have responded to these concerns by providing a wide range of child care programs that better assist their employees to meet their parenting needs with the belief that this will permit their employees to channel their time and energies more fully to work. Not only can employees benefit from such practices but also organizations can profit from these actions by retaining a motivated and productive workforce.

Many large companies provide on-site or near-site child care centers for their employees' children. Employees who avail themselves of these facilities feel more satisfied with their jobs, express greater loyalty toward their companies, and miss fewer days of work. People who are seeking employment are attracted to organizations that offer some type of child care assistance. Organizations, in turn, reap benefits in the form of improved employee productivity and morale. Another way that organizations gain from having these facilities is by being able to form a progressive public image that they can use to attract and recruit better talent.

Some companies are not able to provide on-site child care centers and instead offer alternative arrangements to meet their employees' needs. Such arrangements include, but are not limited to, contracting with individual child care providers and getting discounts for their employees, providing cash vouchers that employees can use at the child care centers of their choice, and offering referral services that help parents with information and advice regarding different child care facilities. Some employers get involved with the communities in which they do business and initiate and/or expand a variety of child care services, such as after-school programs or "sick child care centers." A strategy adopted by many companies involves a revision of their sick leave policies so employees can take sick leave when their family members are ill.

These are some of the more popular and widely used programs that provide direct assistance to individuals with child care needs. However, as previously mentioned, a rapidly growing number of employees are also responsible for the care of their elderly relatives. Employers are increasingly responding to their employees' demands for elder care assistance by providing them with referral services, offering them support groups through their employee assistance programs, and providing social workers who visit their employees' elderly relatives.

Despite the wide range of dependent care programs offered by organizations, a key issue for employees with any kind of dependent care responsibilities involves the freedom to have some degree of control over their time. Organizations offer a variety of flexible work options that allow employees to choose a work schedule that best fits their needs. Some of the widely offered flexible work arrangements are discussed next.

7.2. Flexible Work Arrangements

Flexible work arrangements are designed to help individuals balance their family and work responsibilities by giving them the option of choosing when to work, how much to work, and where to work. One of the earliest and most enduring types of flexible work arrangements—flextime—involves working a core set of hours with flexible starting and leaving times. By giving people more control over their work times, employees are better able to meet their various family obligations. Employees who use flextime report being able to effectively perform their jobs, feeling satisfied with their work, having better relations with their supervisors and team members, and expressing relatively little intention to leave the company. Companies may not only gain in terms of improved employee productivity but also can incur savings through reduced use of sick days, absenteeism, and lateness.

Another option that is increasingly being offered by many companies is that of part-time employment or a reduced-load work arrangement. Companies can allow part-timers to jointly take responsibility for a full-time assignment so that the total job is "shared" by two or more part-timers, an option known as job sharing. Part-time employment was initially proposed as an option for women managers and professionals with young children who often chose to quit their jobs rather than give up time with their children and

continue to work full-time. Often referred to as the "mommy track," the option of part-time employment has been surrounded by controversy with regard to its effectiveness. Some people believe that individuals who opt for part-time work—whether mothers or fathers—perform just as effectively in reduced-load work arrangements as they did when they worked full-time. Moreover, they also believe that employees who participate in such work arrangements feel happier and more satisfied with their work–home balance than they did when they worked full-time.

However, there are others who believe that part-time employment, no matter how brief, can signal to the organization that the employees may not be committed enough to their jobs and organizations. Such stigmatization is more common for women than for men, perhaps because women are much more likely than men to avail themselves of this option or because men are more likely to negotiate informal reduced work involvement with their bosses rather than openly declare their desire to spend more time with their families.

To ensure the success of part-time arrangements, companies have to overcome their stereotypes of people who use these programs and treat them equitably. More fundamentally, companies have to rethink their models of career success. They can offer alternatives to fast-track linear career paths and assignments that can have negative effects on employees' personal and family lives.

Another form of workplace flexibility, telecommuting, focuses not so much on reducing the amount of time employees spend at work but rather on offering employees the option of working from home or from a satellite office for a given portion of the workweek. This allows employees greater discretion in structuring their work and family time. However, working from home may not necessarily solve the employees' child care needs or eliminate work–family conflict. Instead, the blurring of the boundaries between their work and family roles can make individuals increasingly more vulnerable to continuing stressors from both roles. They may also find themselves increasingly isolated from their colleagues and the social interchanges at work that can provide a welcome reprieve from the daily pressures. Nevertheless, telecommuting has much to offer employees who wish to balance their work and family demands.

Another widely used initiative to help employees meet their family obligations is that of family leave. Since the passage of the Family and Medical Leave Act in 1993, a number of companies have instituted various programs that allow their employees to take leave to care for their children, spouse, or parents. Many of these leave polices offer job protection to employees and provide them with medical benefits for a certain period during the leave. By directly addressing employees' needs for caregiving time, family leave programs can improve recruitment and retention of talented employees.

Organizations can offer a patchwork of solutions and programs to enable their employees to fulfill their work and family obligations. However, to be truly effective in the long term, companies have to undergo a fundamental shift and alter their cultures in a way that recognizes and legitimizes work–family issues. In addition, the success of most of these programs depends in large part on the support and commitment of the top management to work–family issues.

In summary, many of these programs may not be able to eliminate work–family interference, but they can provide employees with sufficient flexibility to achieve a more fulfilling balance between their work and family lives.

8. INDIVIDUAL ACTIONS TO MANAGE WORK–FAMILY RELATIONSHIPS

Individuals with extensive work and family responsibilities need to find ways to manage their work–family interface effectively. This involves not only reducing their level of work–family conflict to a reasonable level but also increasing the likelihood that their work and family roles will enrich one another. Given the advantages of working in a family responsive organization discussed previously, individuals seeking employment should assess this critical aspect of a prospective employer's culture and use this information in deciding whether to accept a job offer.

Three proposed principles for creating allies of work and family are worth serious consideration. First, individuals should clarify what is important in their lives. Individuals vary in the relative importance they attach to different parts of their self-identity (e.g., family, work, leisure, and community service), and some people are more aware of their underlying values than others. Self-awareness is crucial for managing the relationship between work and family roles because it helps individuals make effective decisions about the allocation of time and emotion to different roles and the application of resources from one role to enhance the quality of life in the other.

Second, individuals should recognize and support the whole person. This requires a realization of how

participation in different life roles can enrich one's life. It also involves building supportive relationships with individuals in different roles, such as work, family, and community. The support can be tangible assistance (e.g., help around the home) or it can be intangible (e.g., listening to problems and offering advice on solving them). The stress literature indicates that social support from others can enhance well-being by reducing the amount of stress an individual experiences and by helping the individual cope with the stress that is experienced. Seeking support from others requires a willingness to ask for—and accept—others' help. Because of norms of reciprocity and fairness governing human behavior, it also requires a readiness to provide support to others in need of assistance.

Third, individuals should continually experiment with how goals are achieved. This experimentation requires periodically reexamining goals, developing new strategies to accomplish these goals, and adopting new perspectives on one's work and personal life. It involves an openness to change—a willingness to question what is important in life and what approaches are necessary to achieve desired outcomes. Like the problem-focused coping technique lauded in the stress literature, continual experimentation often means extensive communication and problem solving with other people (e.g., spouse and boss) to reduce the conflicts in one's life and to strengthen the quality of one's significant life roles.

9. CONCLUSION

Work and family lives are intertwined in many respects. Experiences at work can influence the quality of family life, and family experiences can affect the quality of work life. Some of these consequences are negative, such as when one role interferes with full and enjoyable participation in the other role. In other cases, however, resources from one role can be used to strengthen the quality of life in the other role. Organizations and individuals can take actions to relieve the stress arising from the intersection of employees' work and family roles and to strengthen the positive effect of each role on the other role.

See Also the Following Articles

Elder Caregiving ■ Family and Culture ■ Leisure and Work, Relationship between ■ Part-Time Work

Further Reading

Barnett, R. C. (1998). Toward a review and reconceptualization of the work/family literature. *Genetic, Social, and General Psychology Monographs*, *124*(2), 125–182.

Barnett, R. C., & Hyde, J. S. (2001). Women, men, work, and family: An expansionist view. *American Psychologist*, *56*, 781–796.

Edwards, J. R., & Rothbard, N. P. (2000). Mechanisms linking work and family: Clarifying the relationship between work and family constructs. *Academy of Management Review*, *25*, 178–199.

Friedman, S. D., & Greenhaus, J. H. (2000). *Work and family—Allies or enemies? What happens when business professionals confront life choices*. New York: Oxford University Press.

Greenhaus, J. H., & Beutell, N. J. (1985). Sources of conflict between work and family roles. *Academy of Management Review*, *10*, 76–88.

Greenhaus, J. H., & Parasuraman, S. (1999). Research on work, family, and gender: Current status and future directions. In G. N. Powell (Ed.), *Handbook of gender and work* (pp. 391–412). Thousand Oaks, CA: Sage.

Haas, L. (1999). Families and work. In M. Sussman, S. K. Steinmetz, & G. W. Peterson (Eds.), *Handbook of marriage and the family* (2nd ed., pp. 571–612). New York: Plenum Press.

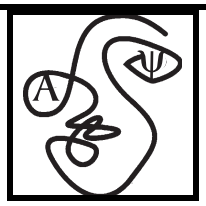
Hall, D. T. (1990). Promoting work/family balance: An organization-change approach. *Organizational Dynamics*, *18*(3), 5–18.

Hochschild, A. (1989). *The second shift: Working parents and the revolution in the home*. New York: Viking.

Lambert, S. J. (1990). Processes linking work and family: A critical review and research agenda. *Human Relations*, *43*, 239–257.

Lobel, S. A. (1991). Allocation of investment in work and family roles: Alternative theories and implications for research. *Academy of Management Review*, *16*, 507–521.

Marks, S. R., & MacDermid, S. M. (1996). Multiple roles and the self: A theory of role balance. *Journal of Marriage and the Family*, *58*, 417–432.



Work Environments

Rotraut Walden
University of Koblenz, Koblenz, Germany

1. Introduction
2. Using Buildings for Increasing Productivity
3. A Historical Perspective on Working Environments
4. Some Recent Studies on Office Environments
5. General Models Describing Environment–Behavior Relations in the Workplace
6. Future Developments
Further Reading

GLOSSARY

Hawthorne effect Effect where if workers notice that they are being observed and that conditions in the workplace change because of them, efficiency increases independently of actual improvements.

personalization The opportunity to make personal changes in the environment through decorations (e.g., photographs, artwork) and in the form of functionality.

postoccupancy evaluation (POE) Evaluation that investigates to what extent a constructed environment provides for current users' needs and demands made by experts.

productivity A synonym for output, efficiency, motivation, individual performance, organizational effectiveness, production, profitability, cost-effectiveness, competitiveness, and/or quality of work of a company or an individual.

sick building syndrome Syndrome where the users of a building become ill as a result of faulty construction and/or through the use of hazardous materials.

user-needs analysis An assessment of the needs, desires, and preferences of future users of a building; it is usually employed during the early phase of the construction program so that potential recommendations can still be incorporated for the improvement of the future building.

Yerkes–Dodson Law of 1908 States that the link between arousal and performance results in an inverse U-shaped curve; performance is better during medium arousal than during low or high arousal.

Performance and well-being are the most common criteria for the improvement of work environments. A glance at history shows the focal points of psychological research over the ages, including advancements in evaluatory methods for improving the working environment. This article presents and discusses some classic studies in this field. Of current interest are building performance evaluation and facility management. The article presents psychological frameworks and models summarizing the current knowledge about environment–behavior relationships in workplaces, envisaging possible amendments to them. Because of the growing importance of information technology, customer service, and globalization, the article concludes by dealing with the foreseeable development of offices and the consequences of internationalization.

1. INTRODUCTION

The aim of environmental psychology of workplaces and organizational psychology is to improve workers' performance and well-being. Performance refers to individual,

©2004 Elsevier Inc.
All rights reserved.

team, or organizational efficiency. Well-being complements the question of satisfaction and is a more holistic approach of recent research that includes health. Satisfaction is generated when, among other things, a worker or an employee can achieve goals and satisfy needs in the workplace. It is an individual's general affective and cognitive evaluation of the job and the physical environment that can support or contrast these goals and needs. The effects of increased performance at reduced costs are of particular interest to companies. It is also important for a business to have creative and innovative individuals who work to meet the constant demands made on the company and who are willing to commit themselves to the company. For users, there is a pivotal interest in the continuity of the organization so as to ensure the stability of their own jobs. Consequently, they concern themselves with the company's performance and make individual contributions to this success. If one considers the proportion of time spent at work over a lifetime, there are compelling reasons for increasing well-being during this period. In the long run, well-being in the workplace is, like salary, an important incentive for employees.

2. USING BUILDINGS FOR INCREASING PRODUCTIVITY

After a review of many different studies, Gifford concluded in 2002 that the majority of companies still do not take into account the fact that even a minimal investment in the workplace can result in an improvement of 10 to 50% in productivity. Companies see their employees as their biggest potential growth factor, but at the same time, employees create the highest costs. A reduction in construction costs usually results in deteriorating working conditions. The strategy for how an investment in new construction or renovation can be implemented is directly connected to the psychological factors of the work environment itself. Therefore, a major question explored in this article is the following: How can performance and well-being in workplace buildings be measured and improved?

3. A HISTORICAL PERSPECTIVE ON WORKING ENVIRONMENTS

Writing about work environments in 2001, Bell and colleagues differentiated among factories, large firms, and offices. This historical retrospective leans closely toward Becker's 1991 article. During the early 1900s, Frederick Taylor assumed that incentives in terms of salary, production techniques, and the work environment should enhance workers' productivity. In fact, it had more to do with the assembly line in the factory. Industrial growth was explosive. Social processes, well-being, and satisfaction—the points of view of the employees—were hardly considered. Scientific management led to routine, standardized jobs and to studies about heating and lighting.

Avoiding Taylorism by using human factors, and an increase in automation of factories to increase efficiency and individual performance between 1910 and 1940, was followed by an emphasis on task performance and human relations between 1940 and 1950. Many results that had been considered valid until that point were questioned in the light of the famous experiments conducted by Elton Mayo and colleagues between 1927 and 1939 at the Western Electric Company plant in Hawthorne, Illinois (near Chicago and Cicero), on the effects of physical work environment on workers' performance. As Roethlisberger and Dickson reported in 1939, work performance increased in both experimental and control groups, although lighting was improved only for the former. In follow-up observations, researchers established that performance improved again despite the fact that illumination had been reduced by 70%. When they replaced a set of lights with bulbs of the same wattage, the employees were even more satisfied. These results gave rise to the following interpretation: If employees notice that the conditions of their immediate environment change because they are being observed, performance increases independently of any actual improvements. This reaction was named the Hawthorne effect. As a result, the naive determinism that dictated that simply improving environmental conditions has a direct effect on desired behavior was questioned. Instead, the role played by the employees' own perceptions, beliefs, preferences, experiences, and personalities was considered more carefully.

Studies on group dynamics, communication, and conflict then followed between 1950 and 1960. These studies were motivated by the need for a decrease in interpersonal conflicts and the development of democratic and cooperative group processes at work. This was often provoked by the preoccupation with World War II. In 1967, Robert Sommer developed a "small group ecology." Arguments for building planning and equipping furnishings, as well as for seating arrangements, formed the origins of environmental psychology. Architects and planners were viewed as managers of human resources.

Subsequently, between 1960 and 1970, a dominant theme was the focus on nonpaying users. Participation of users in the planning of work environments was also encouraged. This should result in a building layout that better mirrors the needs, values, and preferences of its users. Communication, comfort at work, and satisfaction emerged as central concerns in numerous studies carried out between 1970 and 1980.

Studies on office landscape revealed that employees rejected office layouts designed for the ease of communication within the company. There were often complaints about noise and lack of privacy. Workplace research increasingly busied itself with office space, and this corresponded with the increasing value of customer service. For example, in 1984, Brill and colleagues reported the results of an extensive 5-year investigation of 70 firms conducted from the Buffalo Organization for Social and Technological Innovation (BOSTI). The study addressed three basic questions. First, does the overall design of the office have a fundamental influence on employees' productivity and their daily quality of life at work? Second, which specific office characteristics breed these sorts of influences, and in what ways does this happen? Third, how great is the value of these influences in financial terms? Between 1977 and 1983, 6000 people were interviewed over two time periods: before and approximately 8 to 12 months after a generally positive remodeling of their workplace. A wide set of "objective" and "subjective" criteria were used for assessing productivity and office characteristics. After the first investigation, BOSTI recommended targeted changes whose success would be measured after 5 years in terms of monetary value. The results of the BOSTI study suggested that the changes that the firms aimed to make contributed to an increase in efficiency of 15 to 17% in the annual salary of three career groups over the 5-year period: managers, professional or technical workers, and clerical workers. In doing this, investments and the costs of the workplace were calculated.

With regard to work satisfaction, relocation frequency and noise had the highest monetary value. Approximately 25 to 30% of the changes favoring job satisfaction by way of targeted investments contributed to an increase in productivity. Aspects of job satisfaction were more important to clerical workers than to managers and professional/technical workers. Personal storage space for office materials and coats was deemed to be very important by clerical workers. The remaining 70 to 75% of the monetary value was determined by enclosure and layout in job performance.

During the 1980s, Preiser and colleagues introduced the concepts of postoccupancy evaluation (POE) and performance profiles as methods of evaluation for these studies. Today, there are several methods of data collection, such as program development, user needs analysis (UNA), postconstruction evaluation, and POE, that adhere to the comments made by users and experts alike. The POE investigates to what extent a built environment is set into use in accordance with users' needs and the demands made by experts. Therefore, feedback for planners and building owners on the effects of the built environment should be identified, as should the prerequisites for the construction of a better environment. The POE assesses environments according to what extent they encourage their designated function, or at least do not disturb it, and to what extent these environments correspond to the needs, interests, and desires of the users. In comparison with the assessments made by competing architectural firms, the users' view plays a special role.

A broad range of techniques may be used for collecting data, including interviews, questionnaires, checklists, collections of critical events, targeted time-budget studies in the form of log-book entries, objective data taken from thermometers and hygrometers, and counts made from the analysis of physical traces, photographs, and video recordings.

Diversity of methods and the consultation of people in central groups with different visions are a means of choice. The results of the POE are short, medium, and long term. The POE contains feedback from users about problems with the building and proposed solutions. Also, lessons about how to improve the design cycle of a building can be implemented, and planning and design criteria for specific types of buildings can be developed. The immediate goals of these evaluations are to improve building performance with regard to health, security, functionality, and psychological satisfaction. Furthermore, they aim to save costs on maintenance with regard to the whole life spans of buildings. Finally, benchmarks, concepts, and guidelines can be developed for future projects.

During the 1990s, various concepts were also introduced, including facility management, building delivery, and life cycle management. These concepts are also related to the principle of improving building performance and adopting different evaluation methods to combine architects' and users' perspectives on how the improved performance of buildings is perceived. According to a 1997 chapter by Preiser and Schramm, all phases and evaluation methods can be integrated into the building performance evaluation.

4. SOME RECENT STUDIES ON OFFICE ENVIRONMENTS

Currently, more than 50% of people in the United States work in offices. However, throughout history, offices have changed significantly—from the cubicles of the *Ufficis* in Florence, Italy, around 1559 (from which the term "office" was derived), to offices alongside corridors, to open-plan offices using "landscape offices." Group offices have once again become smaller, whereas so-called "combi-offices" are a collection of individual offices grouped around a service and communication core area used by all. The nonterritorial office is equipped with workstations that are not dedicated to any one individual employee but rather are used by lots of different people who share space at different times. This makes sense if employees often have long periods of work outside the office, as in the insurance industry. Hotelling is one form of nonterritorial office for those who do not require a dedicated, personally assigned office. Like a hotel reservation, available offices and appropriate equipment can be reserved. This kind of office lacks opportunities for personalization. It hinders familiarity between colleagues because office neighbors are always changing. In 1997, Schneider and Gentz observed that "intelligent offices" have a floor space of more than 5000 square meters and that their users are various businesses with complex and innovative types of work. These offices are equipped with personal computers, modern conference technology, and access to the Internet, all of which are connected to the individual workstations. The central computers regulate the building, regulate in-house technology, and support facility management. This also provides round-the-clock support for employees, thereby increasing productivity. "Intelligence" is related to the building in its urban context, the total energy needs, and the real estate market. As for the users, intelligence controls changes through the duration of use, avoiding the sick building syndrome.

In 1995, Becker and Steele reported on their study of a new building, compatible with new demands made on a company, that was planned in the state of Michigan. Steelcase is the world's largest manufacturer of contract office furniture. To remain competitive, employees must continually develop new and innovative products. Steelcase's Corporate Development Center (CDC) achieved success through user participation of 450 employees when planning and designing new workplaces (completed in 1989). Top management decided on the site location, building forms, general design concepts, and flow of communication. User focus groups governed the design of individual workstations, shared commons areas, and dedicated project rooms. In so doing, work patterns and environmental requirements were incorporated. The total workplace concept of the CDC includes the idea of integrating decisions such as human resources, information technology, design, construction, building operations, and management. It should be developed as a concept for the entire process in buildings. Informal communication should stimulate creativity. Therefore, teamwork and face-to-face interaction should be encouraged. Different settings were created for different activities. Spatial mobility during the course of daily activities was expected to increase the probability that people would talk about their work informally. Movement through a building should be used to encourage personal contact; therefore, construction did not adhere to the idea of making people go the shortest way while at work.

A second study conducted by the BOSTI was reported in 2001 by Brill and colleagues. Whereas the earlier BOSTI study lacked a specific theoretical model, rudiments of it can be found in this second one. At the very least, in this study a large number of aspects were named and their effects were described. It is hoped, particularly because the results are not yet final, that interconnected conclusions can also be examined. One of the study's authors, Sue Weidemann, observed the following in a personal communication in November 2003: "The BOSTI study reported the final results of the analyses which were based on only the direct predictors of the outcomes. If both direct and indirect predictors were considered, it would be seen that the work environment has an even greater impact than that reported."

During the 15 years between the two BOSTI studies, some long-term and stable business trends have emerged. These trends are called "new officing." The goal of new offices is to use workplaces, technologies, and work processes as an integrated system of enablers to work smarter at every work location. New officing means a radical redesign of workplaces, work from anywhere, and hotelling.

The second BOSTI study examined the effects of new officing through a 6-year investigation involving 13,000 people in 40 business units that was undertaken between 1994 and 2000. Various kinds of evaluative information concerning the effects of the qualities of work settings on employees' work were collected together with ratings of job performance, team performance, and job satisfaction as well as information on employees' learning.

The results suggest that workplaces can most strongly affect job satisfaction, then team performance, and then individual performance. Included among the environmental features with a strong effect on satisfaction and performance are settings supporting distraction-free work, settings for solo work, settings supporting interactions with coworkers, and settings supporting meetings and undistracted group work. To promote these, enclosure is particularly important. Communication is also seen as an important component in satisfaction and performance, and the spatial organization of team workspace can enhance face-to-face communication. The proximity of workstations and shared service supplies can foster informal interaction. Altogether, these results would advocate the use of combi-offices for various purposes. Results also suggest that a carefully designed workplace supporting employees' and teams' activities is an investment that pays off in business terms as well as in positive changes in corporate culture.

5. GENERAL MODELS DESCRIBING ENVIRONMENT–BEHAVIOR RELATIONS IN THE WORKPLACE

Most of the studies described so far lack a general framework or a model that would shed light on the interaction between the physical working environment and the employees' characteristics, behavior, and experiences. In 1986, Eric and Mary Sundstrom developed a framework of the physical environment of offices and factories. In a 1987 chapter, Eric Sundstrom also reported empirical evidence supporting this framework. The analytic framework developed by the Sundstroms treats people and their work environments as interdependent elements of a system and sees environmental relationships as being time dependent. It differentiates among three levels of analysis: individuals, interpersonal relationships, and organizations. At each level of analysis, dynamic processes and different outcomes are involved. The physical work environment may influence various psychological processes such as adaptation, arousal, overload, stress, fatigue, and attitudes. Outcomes for the individual workers can be satisfaction and performance. Changes involve coping as an active attempt to alter any unsatisfactory environment or its effects and other forms of adaptation. Outcomes for interpersonal relationships include adequacy of communication, group formation, and cohesion. An outcome for organizations can be effectiveness. First, the individual worker must deal with ambient conditions such as illumination, color, temperature, air, noise, and music. In addition, the individual's environment includes aspects of the building, its furnishings, its equipment, the individual workstations, and the supporting areas. Second, design aspects such as room layout, building layout, enclosure, and gathering places may influence interpersonal relationships and communication. The regulation of immediacy, choices in communication, and privacy is a key process in these interactions. The office's physical environment, serving a symbolic, self-presentation, or status expression function, can affect job satisfaction. Third, in an organization, people are collected to act in concert toward common goals. In so doing, each person has a specific role in the hierarchy of the company. The facets of physical environment of the organization related to separation and differentiation of work units play a role in this domain. Key processes include congruence of organizational structure and physical environment.

In 2002, Gifford described a model that also tried to cover environmental psychological processes in the workplace in their entirety (Fig. 1). Based on this model, six categories relevant for describing environment–behavior relations at work can be identified. A first category comprises employees' characteristics, including personality, intelligence, cognitive styles, self-concept, emotions, motivation, satisfaction, creativity, skills, experience, training, and job level.

A second category groups the characteristics of the physical environment affecting workers' behavior. The physical environment can be broken down with regard to fixed or shifting, quality of materials, noise, temperature, light, scent, the haptic environment, density of people per space, privacy, territoriality, wayfinding, status, and personalization in displays of self or team. Five fundamental aspects can be filtered out from this: sound (e.g., noise, music), temperature (e.g., heat, cold), air (e.g., pollution, freshness, scent), light and color (e.g., sunlight, incandescent, fluorescent, windows, views), and space (e.g., amount of it, arrangement of work stations). Additions include equipment (e.g., furniture, tools, technology) and displays of self by the users (e.g., improvements in functionality, aesthetic appearance). (According to Walden in 2004, work environments can be assessed using six criteria: functional, aesthetic, social, ecological, organizational, and financial. These criteria are developed by applying the basic central themes of architectural trends, such as "form follows function," to environmental psychology. This means that functional aspects regarding, e.g., layout, wayfinding, and quality of materials, can save time and energy.

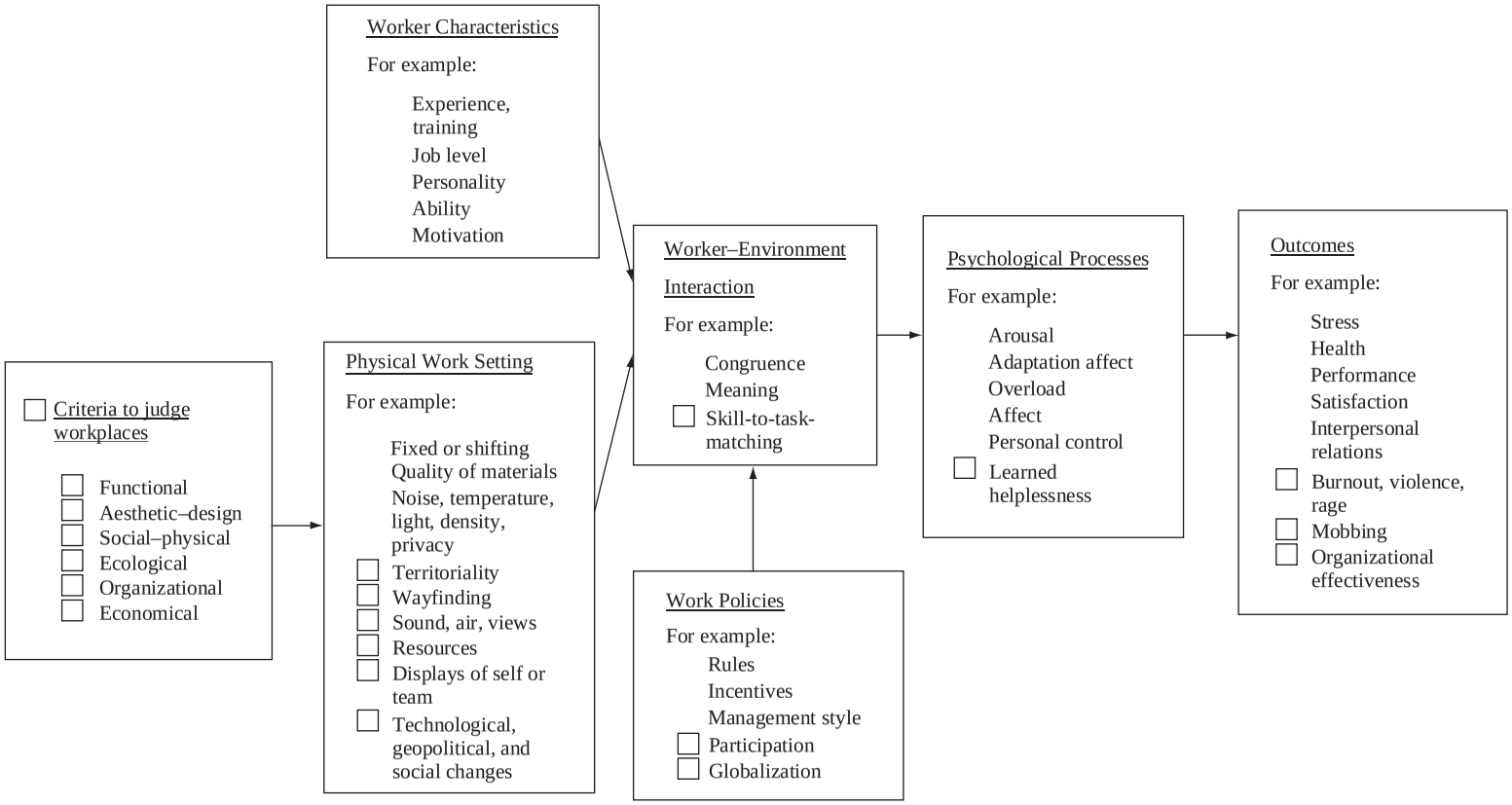


FIGURE 1 A model for workplace environmental psychology. Items designated with empty boxes (□) were added by the author after completion of the literature review. Adapted from Gifford (2002) with the permission of Robert Gifford.

Aesthetic design results in feelings of beauty or newness. Social–physical aspects can result in conflicts that arise from simultaneous use of one setting by multiple parties or in opportunities through communication. Ecological aspects mean that the consequences of a building's existence must be taken into account—from breaking ground, to recycling, to health concerns. Organizational aspects comprise the space–time breakdown of resources, provision of information, materials and storage facilities for logistics, and sharing methods for task cycling. Financial aspects cater to the possibility of the entire organization's purchases and sales as well as its production run.)

A third category refers to work policies.

A fourth aspect is worker–environment interaction in terms of the "fit" between individuals' characteristics and the environment (e.g., skill-to-task matching, the ergonomic fit between humans and machines) as well as the meaning attributed to (and the image of) the working environment. It is important to check whether or not the physical environment and the typical behavioral patterns shown by the user groups (e.g., employees, visitors) match up. The extent to which a fit between a task (especially its level of difficulty) and the abilities of the person involved will be facilitated by the environment should also be considered. Typical examples illustrating the importance of an ergonomic fit between individuals' abilities and perceptions and the use of machines or controllers can be found in control installations for atomic power plants, for space travels, and in the cockpits of airplanes.

A fifth category involves specific individual psychological processes such as arousal, adaptation, overload, personal control (or, oppositely, learned helplessness), and affect. According to the Yerkes–Dodson Law of 1908, performance is better during medium arousal. Higher arousal blocks concentration, and at low arousal a person sleeps. Adaptation occurs when arousal subsides over time due to a process of adjustment. This can also result in a marginal degree of arousal, so that performance is then reduced. Overload arises if too many stimuli emerge over a long period of time, especially when an organism is weakened. (According to Walden, the desire for personal control may result in an appropriation of the environment, that is, in self-regulation of ambient environments and improvements in functionality and aesthetic appearance. Control and user involvement relate to noise, thermal comfort, crowding, space planning, furniture layout, environmental change, and the like. Destructive effects include vandalism and neglect. The latter can be consequences of learned helplessness.)

A sixth category is outcome, defined as the significant changes made to people and to environments through the processes just described. Included among these are stress, health, performance, satisfaction, and interpersonal relations. Morale changes (positive or negative) in the body or the mind are linked to stress and health. (Burnout, violence, and rage can also result from stress. Mobbing is facilitated through loss of privacy.) Performance affects productivity itself but also affects indicators such as resignation rate, time spent on work, and attendance. Feelings relate to satisfaction, evaluation, attitude, emotion, and perception. Interpersonal interaction, work climate, territoriality, privacy, and status are related to social behaviors regarding the way in which space is occupied (i.e., appropriation). The layouts of space, telephones, videoconferencing facilities, and/or e-mail involve components of working conditions and, as such, serve for the transfer of work-related communication. Status demarcation underlines social roles within the hierarchy of a company through objects in the environment. Privacy is optimal at the point between isolation and too many people per room (i.e., crowding). Territoriality marks the possession of space by a single person cut off from others.

Other authors, such as Becker in 1991, have mentioned the improvement of control, communication, environmental change processes, performance, and international influences as the central psychological challenges facing workplaces. Control, communication and performance were already present in the previously described model by Gifford. Environmental change processes, such as office relocation and residential mobility, require adjustment and adaptation. International influences demand increased understanding of the significance of individual organizations and adaptation to labor in other countries.

In 2002, McCoy referred to the multifaceted individual aspects of work environments and emphasized recent empirical findings on the relationship between the physical environment and the health, safety, and behavior of occupants. In addition, she considered aspects related to office development during the past two decades or so. According to her review, more than 50% of working people in the United States work in offices. McCoy mentioned the importance of architectonic details such as decorative styles, personalization, treatment of boundaries, signs, colors, and artwork that encourage a sense of identity and purpose among the individuals in an office. Personalization

includes photographs of family members, hobbies, and/or vacations. These displays of self or team can express the personal commitment to the job and can balance between work and personal activities. The importance of a view for satisfaction and psychological restoration is also emphasized. Access to the necessary resources, the building's furniture and equipment, facilities, and people are part of a system that enables the handling of information for innovations and ordinary office work. Supporting the accessibility of resources can save time for the users and can save money for the organization.

6. FUTURE DEVELOPMENTS

The future developments of workplaces include, on the one hand, the creation of superlative office spaces in the highest buildings in the world such as the Petronas Towers in Kuala Lumpur (Malaysia), a high-rise in earthquake-hit Taipei (Taiwan), and the planning of the Bionic Tower in Shanghai (China). On the other hand, there is an increasing concern for organizational ecology with an emphasis on sustainability. Offices in large complexes are often leased, and conference offices, especially for worldwide teleconferences among members of the same company, are being built at an increasing rate. Linked to this will be hotels with office equipment. Residential offices will have access to neighborhood work centers with telecommuting offices or satellite offices in which large numbers of people can rent well-equipped workstations in their local area or close to other places of activities. Telecommuting includes the virtual office, the home office, and satellite offices. Virtual offices can happen anywhere, including at the client's work site. In these offices, the use of a portable computer and a mobile phone is all that is needed. Trade and factories offer people the chance to work at home. This should help to alleviate the disadvantages of poor public transportation and commuter traffic.

The studies described in this article used multiple methods of investigation to fully understand the complex relationships involved in human behavior and the changing environments of workplaces. In the future, models describing environment–behavior relations at work should be expanded to include the variables that reflect the fast-paced changing conditions of workplaces. More research is needed to understand the value of those physical places and their relationship to their occupants. Even with increasing preconditions for complete freedom of a connection between work and location, there will still be physical workplaces.

See Also the Following Articles

Environmental Design and Planning, Public Participation in ■ Environmental Stress ■ Job Stress ■ Person–Environment Fit ■ Privacy ■ Territoriality

Further Reading

Becker, F. (1991). Workplace planning, design, and management. In E. H. Zube, & G. T. Moore (Eds.), *Advances in environment, behavior, and design* (Vol. 3, pp. 115–151). New York: Plenum.

Becker, F., & Steele, F. (1995). *Workplace by design: Mapping the high-performance workscape*. San Francisco: Jossey–Bass.

Bell, P. A., Greene, T. C., Fisher, J. D., & Baum, A. (2001). *Environmental psychology* (5th ed.). Fort Worth, TX: Harcourt Brace.

Bonnes, M., & Bonaiuto, M. (2002). Environmental psychology: From spatial–physical environment to "sustainable development." In R. B. Bechtel, & A. Churchman (Eds.), *Handbook of environmental psychology* (pp. 28–54). New York: John Wiley.

Brill, M., Margulis, S., & Konar, E., for Buffalo Organization for Social and Technological Innovation. (1984). *Using office design to increase productivity* (Vols. 1–2). Buffalo, NY: Workplace Design and Productivity.

Brill, M., Weidemann, S., Allard, L., Olson, J., & Keable, E. B., for Buffalo Organization for Social and Technological Innovation. (2001). *Disproving widespread myths about workplace design*. Jasper, IN: Kimball International.

Clements-Croome, D. (Ed.). (2000). *Creating the productive workplace*. London: E & FN SPON.

Gifford, R. (2002). *Environmental psychology: Principles and practice* (3rd ed.). Colville, WA: Optimal Books.

McCoy, J. M. (2002). Work environments: The changing workplace. In R. B. Bechtel, & A. Churchman (Eds.), *Handbook of environmental psychology* (pp. 443–460). New York: John Wiley.

McCoy, J. M., & Evans, G. (2004). The physical environment. In J. Barling, E. K. Kelloway, & M. Frone (Eds.), *Handbook of work stress*. Thousand Oaks, CA: Sage.

Preiser, W. F. E., Rabinowitz, H. Z., & White, E. T. (1988). *Post-occupancy evaluation*. New York: Van Nostrand Reinhold.

Preiser, W. F. E., & Schramm, U. (1997). Building performance evaluation. In D. Watson, M. J. Crosbie, & J. H. Callendar (Eds.), *Time saver standards* (7th ed.). New York: McGraw–Hill.

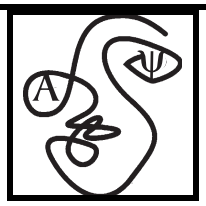
Roethlisberger, F. J., & Dickson, W. J. (1939). *Management and the worker*. Cambridge, MA: Harvard University Press.

Sundstrom, E. (1987). Work environments: Offices and factories. In I. Altman, & D. Stokols (Eds.), *Handbook of environmental psychology* (Vol. 1, pp. 733–782). New York: John Wiley.

Sundstrom, E., & Sundstrom, M. G. (1986). *Work places: The psychology of the physical environment in offices and factories*. New York: Cambridge University Press.

Walden, R. (2004). Assessing the performance of "offices of the future." In W. F. E. Preiser, & J. C. Vischer (Eds.), *Assessing building performance: Methods and case studies*. Oxford, UK: Butterworth–Heinemann.

Wineman, J. D. (1986). *Behavioral issues in office design*. New York: Van Nostrand Reinhold.



Work Motivation

Edwin A. Locke
University of Maryland (Emeritus), College Park, Maryland, USA

1. Needs
2. Values
3. Goals
4. Self-Efficacy
5. Emotions
6. Job Satisfaction
Further Reading

GLOSSARY

emotion The form in which one experiences subconscious value appraisals.
goal The object or aim of an action.
job satisfaction Liking for one's job.
needs The objective requirements of survival and well being.
self-efficacy Task-specific self-confidence.
value That which one considers good or beneficial.

The concept of motivation refers to factors within the individual that impel him or her to action. Motivation is not the same as arousal. Being motivated means being motivated to do something specific. Motivation affects people's choice of action (direction), the amount of effort expended on a given action (intensity), and the persistence of action (duration). In reality, motivation is never separate from cognition. Cognition answers the question, What is? It involves perception and conceptual identification. Motivation pertains to the question, How important is it? It involves evaluation. Evaluation is a form of psychological measurement, with one's values serving as the measuring instrument. For example, cognition might state, Here are three possible courses of action and their possible outcomes. Motivation would state, This one is most important or beneficial to me. Both motivation and cognition are explicitly or implicitly involved in every action.

1. NEEDS

There are five fundamental concepts in the realm of motivation: needs, values, goals, self-efficacy, and emotions. Needs are the objective requirements of an organism's survival and well-being. The existence of needs stems from the fact that life is a conditional process. For a living organism to survive, it must take the actions its nature requires. If an organism's needs are not satisfied, it experiences discomfort, suffering, unhappiness, pain, illness, and/or death, depending on the need in question and the severity and duration of deprivation. Needs are discovered largely through the effects of deprivation. People without food feel hungry and weak; people without self-esteem feel anxious and depressed.

Human needs are most logically divided into two broad categories: physical and psychological. Physical needs are the requirements of a healthy body. They fall within the province of medicine. These needs include air, water, food (including all the nutritional requirements that this implies), rest, sleep, and temperature maintenance. These needs differ in terms of the

©2004 Elsevier Inc.
All rights reserved.

symptoms of deprivation and also in the time lag between total deprivation and illness or death. For example, 2 minutes without air (unless the brain is cooled to a low temperature) leads to irreversible brain damage. A few minutes in very cold water can also be fatal. However, people can survive several days without water and sometimes weeks without food (if they have water). Specific dietary deficiencies (e.g., the B vitamins) may take years to manifest themselves. People may live, and even thrive, for years and even decades with partial deprivation of some physical needs.

Psychological needs are the requirements of a healthy consciousness. The core psychological need is that of self-esteem, the conviction that one is fundamentally worthy and efficacious. The opposite of self-esteem is self-contempt or self-doubt; low self-esteem is often associated with depression. Another psychological need is that of pleasure. There are many possible sources of pleasure in life, including food, drink, work, hobbies, sports, leisure activities, art (e.g., music, literature, and film), friendship, and romantic love (including sex). Art, friendship, and love are also needs in themselves. Art provides humans with spiritual fuel by presenting fundamental value judgments about humans and the world in directly perceivable form.

Contrary to Maslow's well-known need theory (with the following hierarchical order, from low to high: physiological, safety, love, self-esteem, and self-actualization), there is no evidence that needs exist in any built-in hierarchy. Deprivation of some needs causes death faster than deprivation of other needs (e.g., air vs food), but people have the power to prioritize needs (e.g., dive into water and risk drowning in order to save a child). Nor is it the case that all physical needs take precedence over all psychological needs. People may commit suicide in very short order if they conclude that they are totally worthless. Furthermore, people may undergo drastic deprivations or take incredible risks for the sake of preserving a certain view of themselves. In contrast, the deprivation of other psychological needs (e.g., art) may have more gradual effects (e.g., loss of joy in life). It also must be noted that needs are never permanently satisfied; needs operate cyclically. Life is not a one-time, static event but "a process of self-sustaining and self-generated action" (Rand, 1964, p. 16). Some needs may be habitually satisfied (e.g., food), thus leaving the person free from discomfort in relation to that need, but this does not mean that everyone who is satiated with respect to a specific need will automatically proceed to give priority to the same not-yet-fulfilled need.

The concept of need explains why living organisms act at all and must act continually if they are to sustain their lives. The most fundamental reason why people work to earn a living is to fulfill their needs and thus sustain their existence. However, the concept of need alone is not sufficient to explain human action. Need deprivation does not automatically guide people toward the correct action. First, they may not know what to do; pain does not automatically endow one with knowledge of how to relieve it. Feeling hungry does not tell people where to get food or which food to get. Second, people may act against their own needs because they want things that are not good for them, such as marijuana or high-cholesterol foods. Third, there are often options regarding to what need-fulfilling actions people can take (e.g., what career to pursue or which job to apply for). Thus, another motivational concept is required—that of value.

2. VALUES

A value is "that which one acts to gain and/or keep" (Rand, 1964, p. 16). It is what one wants or considers beneficial to one's welfare. Whereas needs are inborn, values are acquired. One needs food but acquires the desire for pizza and ice cream. One needs art but acquires the love of Chopin or Vermeer. Values are a link between needs and action.

Unlike needs, values do exist in a hierarchy, although not an inborn one. If people did not have value hierarchies, they would be paralyzed in the realm of action (not to mention thought). They would not be able to choose among the many action alternatives with which they are constantly faced (e.g., "Should I go to work or stay home and watch TV?") Values prioritize action and therefore prioritize needs.

People are not necessarily fully aware of their value hierarchies. Their subconscious hierarchies may differ from their conscious ones, and this discrepancy may be a source of conflict (e.g., "I want to do well in my work but somehow I feel held back").

Value hierarchies do not remain static. There are many context factors involved, including knowledge, life and career stage, family situation, health, and finances, that can change over time. The following is an important question in the realm of work: How important is one's career in comparison to other values? Highly ambitious people make career achievement a major priority and thus tend to have more career success than those who are less ambitious.

Values also account for individual differences in relation to what career and job one chooses. People differ widely in what they want from their jobs, especially the specific type of work they prefer. Preferences for different types of work are often called interests, but an interest is a form of value (what, specifically, one enjoys doing).

Values are also the basis for the effects of incentives. An incentive is the offer of a value in return for specified actions or results (e.g., a sales increase of 10%). Money is widely used as an incentive because it can be used to purchase many other values and thus is valued itself. People are more motivated at work if they believe that their efforts will lead to a valued reward. This is the basic premise of valence–instrumentality–expectancy theory.

3. GOALS

Values have implications for action but values are general (e.g., "I value education"), whereas action is specific. The last motivational link leading to action is the concept of goal. A goal is the specific object or aim of an action. The word intention is sometimes used as a synonym for goal, although the word intention has more of a short-term focus.

Goals and intentions are the most direct conscious determinants of action. Performance goals are significantly related to actual job performance. The highest level of performance is attained when goals are both difficult and specific, feedback is provided showing goal progress, there is commitment to the goals, and there is knowledge of what strategies are needed to attain the goals.

The motivational sequence, then, goes from needs to values to goals. Goals mediate value effects, and values mediate need effects.

4. SELF-EFFICACY

Although it is discussed elsewhere in this encyclopedia, the concept of self-efficacy must be mentioned here. Self-efficacy is task-specific self-confidence. (Bandura does not use the concept to refer to global efficacy, which, as noted previously, is an aspect of self-esteem.) It is similar in meaning to effort–performance expectancy in expectancy theory, except that it is more inclusive in that efficacy judgments can be based on many factors in addition to effort effects. Although self-efficacy is a belief, it has powerful motivational effects. People with high efficacy set higher goals, are more committed to difficult goals, are more resilient in the face of failure and negative feedback, and choose better task strategies than people with low self-efficacy. There is evidence that goals and self-efficacy mediate the effects of both personality and other incentives, such as assigned goals, money, participation in goal setting, and job enrichment.

5. EMOTIONS

Emotions are the form in which people experience subconscious value judgments. The sequence is as follows: perception (including all stored information about the entity or situation being appraised) → subconscious value appraisal → emotional response. Emotions involve implicit psychological measurements, viz. what is the significance of this in relation to my values? People react differently to the same object when they have different beliefs about it and/or different value standards. Every emotion is the result of a specific form of appraisal. Fear is the response to the appraisal of physical threat, anxiety is the response to a self-esteem threat, anger stems from a perceived injustice, guilt results from the perceived violation of a moral value, etc.

6. JOB SATISFACTION

Job satisfaction stems most directly from the appraisal of one's job as providing important job values or attaining one's goals. (Need-value conflicts can occur when values are irrational—that is, contradictory to one's objective needs.) The most powerful determinant of job satisfaction is mental challenge. This includes being given responsibility for decision making, the chance to use one's abilities, the chance to achieve goals, the chance to learn new things, and feedback regarding one's performance. However, it is important to feel challenged at work one personally likes or considers significant; thus, the work, as noted previously, must be consonant with one's personal interests. Other commonly held job values include competent and considerate leadership and supervision; fair treatment in terms of both content (what rewards one gets, i.e., distributive justice) and process (i.e., procedural justice); the opportunity to advance; competent and compatible coworkers; safe, comfortable, and convenient working conditions; and competitive fringe benefits.

Although perceived job attributes in relation to one's job values are critical determinants of job satisfaction,

other factors are also involved. The individual's core evaluations also play a role. Core evaluations (CEs) are tied to one's needs; they are fundamental estimates that people make of themselves. The four CE components are self-esteem, general self-efficacy, (internal) locus of control, and (low) neuroticism. Core evaluations have been found to affect job satisfaction directly and indirectly through their effects on how people perceive their jobs, the difficulty of the tasks they choose, and the complexity levels of the jobs they hold.

Core evaluations are associated with several of the Big Five personality traits, especially neuroticism (by definition), extraversion, and conscientiousness. Conscientiousness is the best single Big Five predictor of job performance, although its effects are partly mediated by goals. Passion for the work is also important.

The traits of positive and negative affectivity have also been found to be strongly associated with job satisfaction and dissatisfaction. However, it is not clear whether this relationship is causal. Asking people to rate how they feel and then to rate how they feel about their jobs or their lives may simply be asking the same question in slightly different forms.

The issue of the relationship between job satisfaction and job performance has been debated within industrial/organizational psychology for approximately half a century. The correlations between the two tend to be low, although they are higher when corrected for reliability. The correlations are higher for more complex jobs. The real question, however, is the following: What is the causal relationship between them, if any? Certainly, high performance could be a cause of high satisfaction if high performance yielded more inner satisfaction and practical rewards than low performance.

The other side of the coin is more complex. Emotions do contain felt action tendencies; the tendency is to avoid or harm negatively appraised objects and to approach, protect, or retain positively appraised objects. However, a felt action tendency does not compel action and does not identify the many alternative courses of action that one can identify through reason. Fisher and Locke have shown that there are many possible reactions to job dissatisfaction, including job avoidance, aggression (e.g., lawsuits), defiance (of company rules), constructive protest, and inaction, in addition to reducing work effort and productivity. Thus, the question is, What form of action will people choose when they do or do not like their jobs? Given that avoidance is the natural action tendency built into dissatisfaction, one would expect that turnover would be the most likely response to dislike of the job, and, in fact, it is—mediated through the intention to quit and moderated by various factors. It remains to be determined in what circumstances people will choose high or low productivity as their preferred means of expressing their degree of satisfaction with the job. Satisfaction, for example, can lead to setting higher goals or to complacency.

See Also the Following Articles

Emotion ■ Industrial/Organizational Psychology, Overview ■ Motivation and Culture ■ Motives and Goals ■ Values and Culture

Further Reading

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.

Barrick, M., & Mount, M. (2000). Select on conscientiousness and emotional stability. In E. Locke (Ed.), *Handbook of principles of organizational behavior*. Oxford, UK: Blackwell.

Fisher, C., & Locke, E. (1992). The new look in job satisfaction research and theory. In C. Cranny, P. Smith, & E. Stone (Eds.), *Job satisfaction*. New York: Lexington-Macmillan.

Greenberg, J. (2000). Promote procedural justice to enhance acceptance of work outcomes. In E. Locke (Ed.), *Handbook of principles of organizational behavior*. Oxford, UK: Blackwell.

Hackman, R., & Oldham, G. (1980). *Work re-design*. Reading, MA: Addison-Wesley.

Howard, A., & Bray, D. (1988). *Managerial lives in transition*. New York: Guilford.

Judge, T. (2000). Promote job satisfaction through mental challenge. In E. Locke (Ed.), *Handbook of principles of organizational behavior*. Oxford, UK: Blackwell.

Judge, T., Locke, E., & Durham, C. (1997). The dispositional causes of job satisfaction: A core evaluations approach. In L. Cummings, & B. Staw (Eds.), *Research in organizational behavior* (Vol. 19). Greenwich, CT: JAI.

Judge, T., Locke, E., Durham, C., & Kluger, A. (1998). Dispositional effects on life and job satisfaction: The role of core evaluations. *Journal of Applied Psychology*, *83*, 17–34.

Judge, T., Thoreson, C., Bono, J., & Patton, G. (2001). The job satisfaction–job performance relationship: A qualitative and quantitative review. *Psychological Bulletin*, *127*, 376–407.

Lazarus, R., & Lazarus, B. (1994). *Passion and reason*. New York: Oxford University Press.

Locke, E. (1976). The nature and causes of job satisfaction. In M. Dunnette (Ed.), *Handbook of industrial & organizational psychology*. Chicago: Rand McNally.

Locke, E. (2000a). Motivation, cognition and action: An analysis of studies of task goals and knowledge. *Applied Psychology: An International Review*, *49*, 837–853.

Locke, E. (2000b). *The prime movers: Traits of the great wealth creators*. New York: AMCOM.

Locke, E. (2001). Self-set goals and self-efficacy as mediators of incentives and personality. In M. Erez, U. Kleinbeck, & H. Thierry (Eds.), *Work motivation in the context of a globalizing economy*. Mahwah, NJ: Erlbaum.

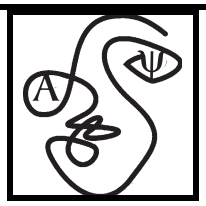
Locke, E., & Latham, G. (1990). *Goal setting and task performance*. Englewood Cliffs, NJ: Prentice Hall.

Locke, E., McClear, K., & Knight, D. (1996). Self-esteem and work. In C. Cooper, & I. Robertson (Eds.), *International review of industrial & organizational psychology*. Chichester, UK: Wiley.

Peikoff, L. (2001). *Objectivism: The philosophy of Ayn Rand*. New York: Dutton.

Rand, A. (1964). *The virtue of selfishness*. New York: Signet.

Rand, A. (1969). *The romantic manifesto*. New York: World.



Work Role, Values Sought in the

Donald G. Zytowski

National Career Assessment Services, Inc., Ames, Iowa, USA

1. General Values
2. Work Values
3. Conceptual Issues
4. Assessment Issues
5. Validity
6. Case Illustration
7. Conclusion
   Further Reading

GLOSSARY

industrial/organizational psychology The branch of psychology that is concerned with worker performance and satisfaction from the perspective of the employer. It emphasizes worker selection and placement and also job descriptions in the service of employers.

job satisfaction The hedonic tone, or the level of satisfaction or dissatisfaction, associated with one's job or occupation. It is assessed in either of two modes: a general statement in scalar response form or responses to a number of facets of work, such as income, work conditions, and prestige. These facets are typically the conceptual source of work values.

vocational psychology The branch of psychology that is concerned with work and careers as they apply to the individual. It emphasizes young people's career development and planning and adults' adjustment to their careers, maximizing satisfactory vocational adjustment.

Work values are personal preferences for selected rewards of work or of a job. They are one of several individual difference variables in the affective domain that are involved in the career decision making and planning of youth as well as in the work behaviors and job satisfaction of employed persons.

1. GENERAL VALUES

A value may be generally defined as that which a person believes is good, true, or beautiful. There are many classes of things that may be valued. Rokeach specifies two classes: instrumental (ways of behaving) and terminal(end states of existence). Others distinguish personal values such as respect for others, belief in God, or living frugally; moral values such as honesty; and cultural values such as the importance of family. Different cultures or individuals may value opposites, as in valuing military service versus pacifism.

2. WORK VALUES

The concept of work values emerged from the study of job satisfaction. It was found that generalized expressions of job satisfaction were approximately equivalent to the cumulated satisfaction with specific aspects of jobs, such as income, hours worked, opportunities for advancement, and level of independence. From inquiring about how satisfied one is with, for example, the work aspect of income, it follows to inquire as to how important (how valued or how needed) income is to the individual.

©2004 Elsevier Inc.
All rights reserved.

Assessment of the importance of various aspects of work and jobs forms the basis for the measurement of work values. A number of work values inventories have been published for both theoretical investigations and use in employee selection and career development.

Work values reside securely in the affective domain of behavior, somewhat less specific than interests (preferences for activities, although sometimes occupations and school subjects) but more specific than personality characteristics. Super proposed an "onion model" with personality at the core and values and then interests in successive layers, with increasing responsiveness to external, environmental factors.

3. CONCEPTUAL ISSUES

The concept of work values suffers from several issues that are yet to be resolved. Most prominent are definitions of terms, their number, and their underlying structure.

3.1. Definitions

An unfortunate conceptual confusion underlies our comprehension of work values. Dawis and Lofquist posit needs in the individual and reinforcers in work. For them, values are the underlying dimensions of needs. Super, on the other hand, specifies values as objectives that one seeks in order to satisfy more basic needs. The definition given previously attempts to avoid the need–value problem by offering preferences in their place and the rewards of work, jobs, or occupations as the objects of those preferences.

3.2. The Number of Work Values

A persistent question concerns the optimal number of work values. Dawis and Lofquist list 21, which have been found to be distributed among six higher order factors. Super first issued a measure of 15 work values similar to those of Dawis and Lofquist and then a subsequent version that added 6 more. (This latter edition has been used extensively in cross-cultural studies.) Other assessments in published inventories and on-line career exploration programs contain as many as 30 work values. The potential number may be nearly infinite, limited only by the eventual triviality of the aspects preferred (e.g., cubicle or window office and reserved parking privileges).

Another complicating factor in delineating the nature and extent of work values as aspects of jobs or occupations is that some job attributes can be bipolar in valence. Danger in an occupation (such as firefighting) can be attractive to some but repellant to others.

3.3. Underlying Structure

Some investigators have attempted to develop the structure of work values. A popular ad hoc structure assumes three sources of satisfactions: the work (solving problems, recording data, etc.), concomitants or work situation (supervision, independence, and work environment), and outcomes (pay and prestige). These are not unlike the classification of the sources of job satisfaction into categories of the work and the conditions of work. A factor analysis by Dawis and Lofquist of their 21 needs revealed six defining scales that Dawis speculates are bipolar and constitute the dimensions of work values: Achievement vs Comfort, Altruism vs Status, and Autonomy vs Safety. Other schemes classify work values as Motivation and Hygiene factors; as Social, Power, Uncertainty, and Goal orientations; or Careerist, Wage Earner, and Entrepreneur.

4. ASSESSMENT ISSUES

A number of measures of work values are available. They are in part distinguished from one another by their method of measurement. One method is Rokeach's pair comparison; the other is the scaled, or Likert, response format. Pair comparison yields a rank order of values without a certain zero point. This is a handicap to statistical analysis but is eminently practical in vocational and career counseling. Also, as the number of variables compared increases, the number of items in a pair comparison inventory increases exponentially, which respondents are apt to view as tedious. Assessments employing scaled responses typically tap need or importance level rather than preference, and they depend for their interpretation on norms.

In general, assessment of work values has not attained the level of sophistication typical of measures of vocational interest.

5. VALIDITY

The stability of work values has been assumed, perhaps from the findings for interests, but there is little empirical evidence in support of this assumption. Values apparently are little related to abilities, somewhat more

related to interests, and more related to personality traits, with occasionally as much as a .40 correlation.

Occupations are far less homogeneous with respect to work values than to interests; thus, work values do not differentiate among occupations as effectively as interests. Patterns of certain work values may be more associated with occupational levels, such as professional compared with skilled trades. Differences are reliably observed in the work values of people of diverse national or cultural origins who are employed in the same job or work assignment by multinational employers.

Given the genesis of work values in the job satisfaction literature, it would be expected that they would predict job satisfaction. Dawis and Loquist demonstrate this very convincingly. Additionally, work values are better predictors of job satisfaction than interests and better predictors of job dissatisfaction than satisfaction. Note that the criterion variable is job satisfaction, not occupational satisfaction. Moreover, work values have been linked with job tenure and some performance variables (sales success), and even in one study with the incidence of coronary heart disease.

6. CASE ILLUSTRATION

"Jenny" is a 30-year-old woman currently employed as a legal assistant. She grew up in a rural environment, living across a gravel road from her grandparents, but now lives and works in the largest metropolitan area in her state. She says she likes her job, largely because her employers are frequently out of the office, leaving her to work independently. She is satisfied with her pay but finds the work sometimes lacks challenge. At home, she has a substantial collection of house plants, and she could imagine working in a plant nursery, if not owning one, someday.

Her profile on a revised version of Super's Work Values Inventory reveals highest (top quartile) ranks on Security and Lifestyle (being able to do what you want after work). Independence is below the 75th percentile level but more separated from the next lowest score than from the highest two. In the bottom quartile are Mental Challenge, Creativity, Achievement, and Variety. Between these extremes are Income, Coworkers, Supervision, Work Environment, and Prestige.

It is not good practice to attempt to reconcile a work values profile exactly with an individual's current employment. Other job factors (e.g., the employment market) and other values (e.g., family matters) may operate. Jenny's current position apparently satisfies a sufficient number of her work values for her to continue in it. Jenny might use her results to refine her current position, to consider a next position, or possibly to pursue the horticulture field. Although her value on Lifestyle and Independence is not inconsistent with owning a plant nursery, she might do well to consider how she might achieve this in light of her highest rank on Security and her very low score on Achievement. One would not use the results of her Work Values Inventory to validate or caution against Jenny's hopes but, rather, to help her envision how she might attain them and function within them should she succeed.

7. CONCLUSION

Despite their relatively long history in differential psychology, the concept of work values appears to be poorly conceptualized. Several measures of work values have been developed and have demonstrated some utility. They appear to account for significant variance in job satisfaction and, to a lesser extent, job performance. Work values do not apparently differentiate occupations, but they have utility in assisting vocational development and career planning of individuals.

See Also the Following Articles

Industrial/Organizational Psychology across Cultures
- Industrial/Organizational Psychology, Overview
- Vocational Psychology, Overview

Further Reading

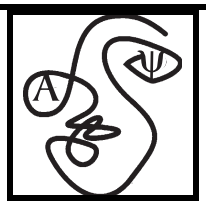
Dawis, R. V. (1991). Vocational interests, values and preferences. In M. D. Dunnette, & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (2nd ed., Vol. 2, pp. 833–872). Palo Alto, CA: Consulting Psychologist Press.

Dawis, R. V., & Lofquist, L. H. (1984). *A psychological theory of work adjustment*. Minneapolis, MN: University of Minnesota Press.

Hofstede, G. H. (2001). *Culture's consequences: Comparing values, behaviors, institutions and organizations across nations*. Thousand Oaks, CA: Sage.

Rokeach, M. (1973). *The nature of human values*. New York: Free Press.

Sverko, B., & Super, D. E. (Eds.). (1995). *Life roles, values, and careers*. Palo Alto, CA: Davies-Black.



Work Safety

Dov Zohar

Technion–Israel Institute of Technology, Haifa, Israel, and Institute for Work and Health, Toronto, Canada

1. Individual-Level Perspective
2. Group-Level Perspective
3. Organization-Level Perspective
Further Reading

GLOSSARY

accident proneness An invalidated hypothesis suggesting that some individuals are frequently involved in accidents due to accident-prone personality.

behavior-based safety An intervention approach using Skinner's behavior modification principles (frequent contingent reinforcement) to overcome tendencies promoting unsafe behavior.

developmental hypothesis A proposition suggesting that young workers suffer heightened injury due to factors associated with lack of emotional maturity, adventure seeking, underestimating risks, and carelessness at young age.

healthy organization An organization that actively promotes employees' physical and emotional welfare.

human factors engineering An interdisciplinary field of research focusing on human and machine (or job) attributes influencing the degree of fit between the two.

inexperience hypothesis A proposition ascribing heightened injury of young workers to limited knowledge and skills as well as little or no previous work experience; that is, age covaries with experience.

melioration effect A strong tendency to overweight immediate or short-term outcomes, even when this compromises expected utility in the long run; melioration promotes unsafe behavior, whose short-term benefits often include greater speed, lesser effort, and personal comfort.

recency effect A strong tendency to overweight recent or frequently recurring outcomes, even when this compromises expected utility in the long run; recency and melioration exert additive effects on unsafe behavior.

safety climate A consensual perception of employees concerning the priority of safety in the workplace, derived from enacted policies and practices indicative of managerial commitments.

supervision-based safety An intervention approach using readily available supervisory incentives (recognition and feedback) to change the unfavorable cost/benefit ratio associated with safe behavior.

transformational leadership A value-based (vs transaction-based) leadership pattern characterized by commitment to the psychological and physical welfare of group members.

Work accidents cost the American economy an estimated $132 billion per year in addition to human suffering and loss of life. A fatal work injury occurs every 2 hours, and a disabling injury occurs every 8 seconds. Overall world estimates suggest that approximately 100 million work injuries, 100,000 of which are fatal, occur each year. Because 85 to 90% of accidents involve a combination of unsafe acts and unsafe conditions, it is evident that human behavior is at the core of occupational safety. This suggests a significant role for psychology in terms of explanatory models and alleviation practices. Theoretical explanations and corresponding interventions can be grouped by level of analysis, stemming from the fact that behavior in organizations is influenced by individual-, group-, and

©2004 Elsevier Inc.
All rights reserved.

organization-level factors and that a comprehensive description requires cross-level integration. This review of significant contributions of psychology to work safety is organized accordingly.

1. INDIVIDUAL-LEVEL PERSPECTIVE

Individual-level explanations of unsafe behavior focus on pathogenic factors, mostly personality traits, youthfulness/inexperience, and substance abuse. Recent developments offer complementary (nonpathogenic) explanations focusing on known behavior biases that are likely to promote unsafe behavior during routine work.

Early research on behavioral safety was concerned with the accident proneness hypothesis, stimulated by studies during the 1920s that indicated that some individuals are frequently involved in accidents. Interest in the accident-prone personality waned due to researchers' inability to establish its core characteristics and to indications that proneness was an artifact of inadequate statistical analysis. Subsequent studies focused on individual personality traits and their relationship with accident involvement. This work is largely empirical, consisting of weak arguments and post hoc explanations. Literature reviews concluded that extraversion (sensation/excitement seeking) and neuroticism were associated with increased accident likelihood. Subsequent reviews added aggression, social maladjustment, impulsivity, and external locus of control to extraversion and neuroticism as relevant personality attributes. More recently, researchers tested relationships between the five primary personality dimensions and accidents, reporting a relationship among neuroticism, low conscientiousness, low agreeableness, and openness. However, because effect sizes have been small to moderate and no coherent theoretical framework has emerged (i.e., each personality dimension requires a different post hoc explanation), the utility of personality testing as a screening device for high-risk jobs remains questionable.

The second pathogenic factor is known as "young workers." Adolescents and young adults sustain injuries at two to four times the rate of those over 25 years of age. Research focused on two rival hypotheses. The developmental hypothesis ascribes heightened injury to factors associated with lack of emotional maturity and cognitive development, greater adventure seeking, limit testing, underestimating risks, and carelessness at young age. The inexperience hypothesis ascribes heightened injury to limited knowledge and skills as well as to little or no previous work experience (i.e., age covaries with experience). Empirical evidence favors the inexperience hypothesis, suggesting that the injury rate in all age groups is highest immediately after starting a new job (Fig. 1). Therefore, the lack of experience associated with young age and overrepresentation of young workers in short-tenure jobs involving frequent job changes are the underlying factors that drive elevated injury rates for young workers. This finding has important practical implications because inexperience, unlike developmental factors, can be addressed through better training and personal coaching. These data emphasize the importance of early safety training and supervised entry given that the first few weeks at a new job are critical.

Substance abuse (mostly alcohol) is the third pathogenic factor given the prevalence of medium to heavy drinking in blue-collar workers. Literature reviews reveal mixed or inconsistent evidence, mainly due to methodological shortcomings of available research. Most studies fail to discriminate between on-the-job drinking and off-the-job drinking because respondents were only asked to report substance use "during the past month." However, studies that included this distinction found on-the-job substance use to be a significant predictor of work injuries. This agrees with laboratory research that revealed impairment of basic cognitive and psychomotor skills, for example, impaired time estimation, divided attention, vigilance, and complex reaction time after consumption. Therefore, workplaces where alcoholic drinks are allowed/available (e.g., during lunch breaks) run a greater risk of alcohol-related accidents. These effects indicate a need for screening of alcoholic individuals, employee assistance programs, and policy implementation regarding on-site consumption. However, substance use may also be related to increased risk of injury due to a common cause such as engagement in deviant behavior, general disobedience, or impulsive behavior. Job stress may also raise the likelihood of both substance abuse and accidents.

Decision-making research provides alternative explanations of unsafe behavior (i.e., without reverting to pathogenic factors), drawing on universal behavioral biases that can also explain unsafe behavior in routine work. A cost/benefit analysis incorporating such biases suggests that appraised benefits of unsafe behavior often outweigh those of safe behavior. One relevant bias, known as the melioration effect, stipulates that decision makers tend to overweight immediate or short-term

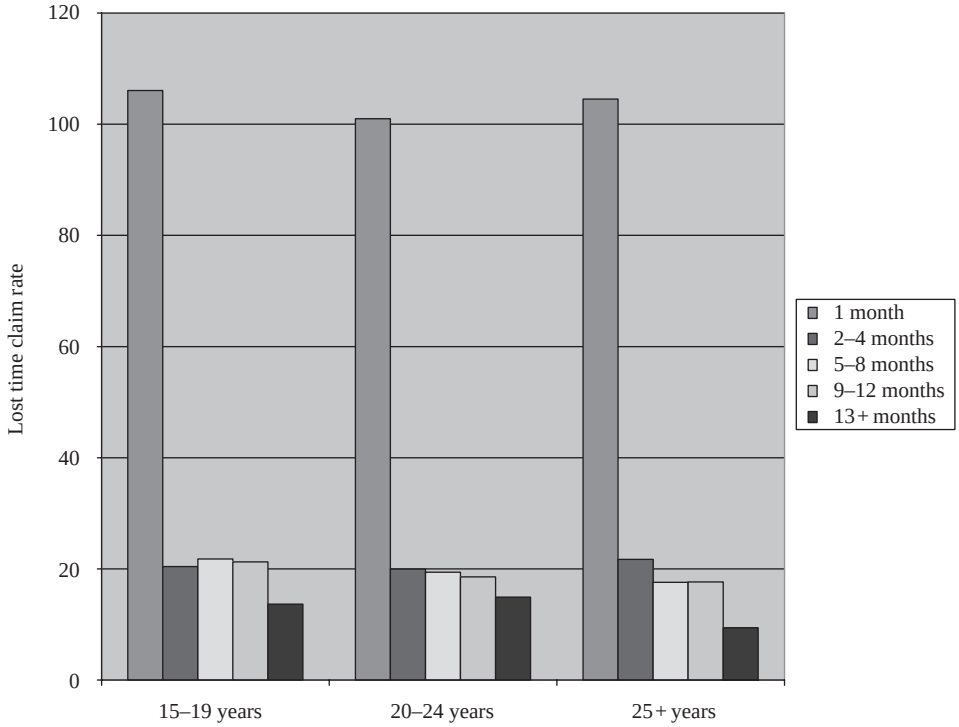


FIGURE 1 Graph illustrating that first-month lost time claim rates are higher at any age. Reprinted from Breslin and Smith (2003).

outcomes, even when this compromises expected utility in the long run. Consequently, the immediate cost of safe behavior, such as slower pace, extra effort, and/or personal discomfort, is often given greater weight than are low-probability and longer term risks avoidable by safe behavior. A second bias is the recency effect, which stipulates that decision makers tend to overweight recent outcomes when making decisions, even when this may compromise expected utility in the long run. This tendency results in underweighting low-probability outcomes, even when they may carry large costs (e.g., injury). Because unsafe behavior leads to immediate and frequent benefits (e.g., increased pace, greater comfort) and safe behavior results in nonevents at best (i.e., avoidance of low-probability injury), the "benefits" of unsafe behavior are likely to be overweighted. Jointly, these biases imply that choosing safe behavior in routine work is not as natural as one might assume based on the self-preservation postulate (i.e., physiological needs as the most basic needs).

1.1. Intervention Approaches

The best individual-level interventions are based on the complementary paradigms of human factors engineering and behavior modification. Human factors/ergonomics interventions are based on the premise that errors and accidents arise from various types of person–job misfits where job demands exceed operators' (mental or physical) capacity. Using job-analytic techniques, key misfits are identified, leading to job redesign and better human–machine interface where the likelihood of error or overload is reduced. Behavior modification approaches combine safety training and behavior modification. The effect size nearly doubles when training is combined with behavior modification in comparison with training alone. Safety behavior modification, also known as behavior-based safety (BBS), is an industrial/organizational application of clinical and educational behavior modification programs stemming from Skinner's reinforcement principles. BBS interventions modify the subjective expected utility or valence of safe behavior by offering rewards whose frequency and immediacy enlist the melioration and recency effects in favor of safe behavior. BBS interventions typically employ publicly displayed charts of the frequency (percentage) of focal safety behaviors in comparison with designated targets. Data are based on daily or weekly behavior sampling observations by external observers or coworkers, and

the frequently updated discrepancy between the current level and the designated goal provides the necessary incentive or reinforcement for change (i.e., discrepancy reduction signals future rewards). Although BBS interventions have proven to be effective, maintaining change after program termination is still problematic (i.e., rapid extinction).

2. GROUP-LEVEL PERSPECTIVE

The main factors known to affect group safety performance are leadership and intragroup processes. Leadership is among the best predictors of safety behavior. According to the full-range leadership model, this relationship stems from the fact that effective transactional leaders organize the work better and get subordinates to perform more reliably through better monitoring and rewarding practices. Transformational leaders facilitate open communication, promote learning and development, and encourage workers to shift from passive compliance to active participation—all of which have been shown to promote safety performance. Transformational leaders are also concerned with safety issues because their interactions with group members are characterized by balanced social exchange, nurturing mutual commitment to psychological and physical welfare. This stands in contrast to transactional leaders whose safety commitments depend on external circumstances, mainly whether senior managers consider safety to be an important performance indicator.

Intragroup processes have been the subject of research since Trist and colleagues' seminal work on sociotechnical systems during the early 1960s. This work tested different work designs in coal mines, including autonomous work groups in which group members performed multiple tasks and share the responsibility for meeting production goals. Results indicated improved productivity and safety, apparently due to better coordination among group members due to better understanding of each other's work and higher motivation/alertness stemming from more decision latitude and less monotony. A recent literature review of autonomous teams highlighted inconsistent results yet suggested an overall positive effect on safety outcomes stemming from more effective group processes such as coordination of action around interdependent tasks and mutual concern for members' well-being associated with greater group cohesion.

Intragroup processes have also been studied from a cognitive perspective, focusing on teams performing tasks that demand high performance reliability (e.g., flying aircraft, manning control rooms in chemical installations). Effective teams develop common mental models (i.e., shared understandings of their task/team system) that provide consensual expectations and anticipation of other team members' actions. Perceptual convergence allows team members to synchronize and coordinate task-related interactions in addition to monitoring and providing backup when needed. This results in fewer errors and, hence, increased safety. Although most of the work on mental models has been done with teams performing high-reliability tasks that involve periods of peak demand (e.g., cockpit crews), these results can be generalized to ordinary teams.

2.1. Intervention Approaches

The close connection between leadership and safety indicates that leadership training is a viable approach to improved safety performance. This can be done through reputable leadership development programs (i.e., pending availability of objective evaluation data). Safety-oriented leadership training is also effective. The most extensively tested training program employs a transactional leadership perspective designed to augment supervisory safety-oriented monitoring and rewarding practices (known as supervision-based safety). The results suggest that when supervisors actively monitor and reward safety performance, subordinates modify their behavior accordingly, resulting in effect size equaling or surpassing that of behavior-based safety interventions (Fig. 2). Furthermore, if top management is included in the intervention and supervisory roles are redefined to include safety performance as an important criterion, the effect of the intervention does not diminish after it has ended.

Intragroup processes must be assessed in terms of the distinction between action processes and interpersonal processes. Action processes involve collective monitoring, coordination, and backup, deriving from shared mental models and mutual understanding of the task/team system. It follows that process improvement must involve cross-training and continuous learning as main vehicles for shared understanding. Learning must be integrated in daily routines rather than restricted to formal training (and retraining) sessions, using routine mistakes and successes as learning opportunities (e.g., mission/work shift debriefings). Interpersonal processes, on the other hand, concern conflict management, collective confidence/efficacy, and team cohesion. The most popular intervention approach here is team building, often allied with organization development interventions.

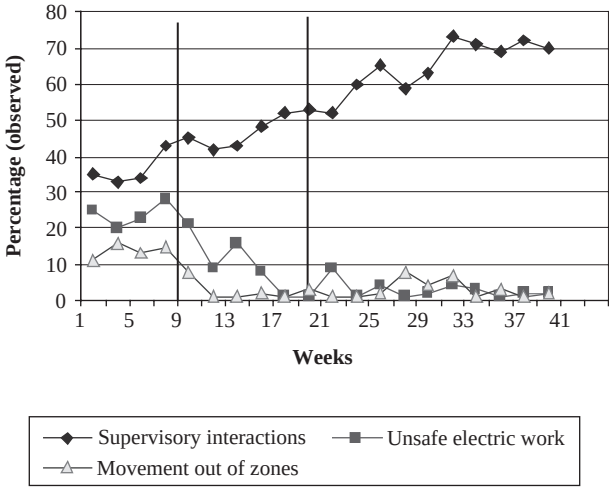


FIGURE 2 Synchronous change in frequency of safety-oriented supervisory interactions and workers' unsafe behaviors during a 3-month supervision-based safety intervention. Reprinted from Zohar and Luria (2003).

Team-building programs involve introducing participative goal setting, improving interpersonal relations based on openness and sharing, working through major problems in the team, and clarifying roles of individual team members. Despite the popularity of organization development, literature reviews have found only minor effects on group performance criteria, including errors and safety. This has led to the development of better balanced approaches combining elements from both processes. The most widely used approach is crew resource management (CRM), which uses detailed performance feedback to create shared mental models and social skills. Although CRM was developed for aviation crews, evaluation studies suggest its applicability to other industrial contexts.

3. ORGANIZATION-LEVEL PERSPECTIVE

Organization-level explanations for unsafe behavior in routine work include safety climate as the main construct. Safety climate emerges from employees' individual assessments of managerial concern for their safety and health. Such concern is expressed and operationalized by managerial practices in routine and nonroutine work conditions. When a group of employees share the same assessment and agree on the priority of safety in the workplace, individual perceptions become organizational climate. Thus, safety climate perceptions refer to attributes of managerial actions that indicate the priority of safety in a company or the importance of acting safely while performing a job. This allows employees to assess the probable consequences of working safely or fast. Climate perceptions clarify supervisory expectations, resulting in a positive relationship between safety climate level and workers' safety behavior, as has been demonstrated repeatedly in relevant studies. Research evidence also offers consistent evidence of a positive relationship between transformational leadership and safety climate stemming from greater concern for members' welfare. Although the terms "safety climate" and "safety culture" are used interchangeably in the relevant literature, only the former is accurate considering the nature of measurement scales involved.

The other main organization-level explanation involves human resources management (HRM) practices in "healthy" organizations (i.e., organizations that actively promote employees' physical and emotional welfare). Practices in healthy organizations reflect employees' commitment (rather than managerial control) as the primary managerial orientation. As in autonomous work group designs, commitment-based orientation presupposes that workers will perform better (and more safely) if decision making is decentralized and is complemented by training, information sharing, compensation adjusted to skill and responsibility levels, and employment security. Furthermore, in the spirit of transformational leadership, HRM practices in healthy organizations are based on mutual concern for physical and emotional well-being, so that line managers are directly responsible for workers' safety, in contrast to the control orientation where injuries are attributed to workers' behavior and noncompliance with safety procedures. Workers reciprocate with proactive safety behaviors (also known as safety citizenship behavior) such as helping coworkers, communicating safety information, and initiating safety improvements. These HRM practices have also been identified as "best practices" in high-performance organizations, identifying a common denominator between organizational effectiveness and safety performance.

3.1. Intervention Approaches

The climate and HRM perspectives suggest that managerial safety practices must change before safety behavior is improved. This creates a link between organization-level leadership and safety performance. Modifying HRM practices toward a commitment-based

orientation requires a transformational leadership perspective. This premise is confirmed by HRM practices in high-performance organizations. Because climate originates from enacted policies and practices (rather than from formal policy statements), improved safety climate requires greater managerial involvement throughout the organizational hierarchy.

See Also the Following Articles

Engineering Psychology ■ Safety and Risks: Errors and Accidents in Different Occupations

Further Reading

Bass, B. M. (1990). *Bass & Stogdill's handbook of leadership*. New York: Free Press.

Breslin, F. C., & Smith, P. (2003). *The relationship between job tenure and lost-time claim rates among adolescent, young adult, and adult workers*. Montreal: Canadian Association of Research on Work and Health.

Cellier, J. M., Eyrolle, H., & Bertrand, A. (1995). Effects of age and level of work experience on occurrence of accidents. *Perceptual and Motor Skills*, *80*, 931–940.

Frone, M. R. (1998). Predictors of work injuries among employed adolescents. *Journal of Applied Psychology*, *83*, 565–576.

Frone, M. R. (1999). Work stress and alcohol use. *Alcohol Research and Health*, *23*, 284–291.

Geller, E. S. (1996). *The psychology of safety*. Randor, PA: Chilton Book.

Hansen, C. P. (1988). Personality characteristics of the accident involved employee. *Journal of Business and Psychology*, *2*, 346–365.

Hofmann, D. A., & Morgeson, F. P. (2003). The role of leadership in safety. In J. Barling, & M. Frone (Eds.), *The psychology of workplace safety* (pp. 159–180). Washington, DC: American Psychological Association.

Keehn, J. D. (1961). Accident tendency, avoidance learning, and perceptual defense. *Australian Journal of Psychology*, *13*, 157–169.

Macdonald, S. (1997). Workplace alcohol and drug testing: A review of the scientific evidence. *Drug and Alcohol Review*, *16*, 251–259.

National Research Council. (1998). *Protecting youth at work*. Washington, DC: National Academy Press.

Pfeffer, J. (1998). Seven practices of successful organizations. *California Management Review*, *40*, 96–124.

Salgado, J. F. (2002). The Big Five personality dimensions and counterproductive behaviors. *International Journal of Selection and Assessment*, *10*, 117–125.

Stajkovic, A. D., & Luthans, F. (1997). A meta-analysis of the effects of organizational behavior modification on task performance, 1975–95. *Academy of Management Journal*, *40*, 1122–1149.

Turner, N., & Parker, S. K. (2003). The effect of teamwork on safety processes and outcomes. In J. Barling, & M. Frone (Eds.), *The psychology of workplace safety* (pp. 35–62). Washington, DC: American Psychological Association.

U.S. Bureau of Labor Statistics. (2002). *Workplace injuries and illnesses in 2001*. Washington, DC: U.S. Department of Labor.

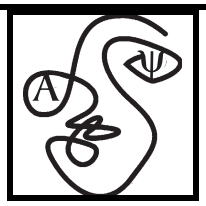
Weick, K. E. (1995). *Sensemaking in organizations*. Thousand Oaks, CA: Sage.

Zacharatos, A., & Barling, J. (2003). High-performance work systems and occupational safety. In J. Barling, & M. Frone (Eds.), *The psychology of workplace safety* (pp. 203–222). Washington, DC: American Psychological Association.

Zohar, D. (2000). A group-level model of safety climate: Testing the effect of group climate on micro-accidents in manufacturing jobs. *Journal of Applied Psychology*, *85*, 587–596.

Zohar, D. (2003). The influence of leadership and climate on occupational health and safety. In D. Hofmann, & L. Tetrick (Eds.), *Health and safety in organizations* (pp. 201–230). San Francisco: Jossey–Bass.

Zohar, D., & Luria, G. (2003). The use of supervisory practices as leverage to improve safety behavior: A cross-level intervention model. *Journal of Safety Research*, *34*, 567–577.



Work Teams

Steve W. J. Kozlowski
Michigan State University, East Lansing, Michigan, USA

Bradford S. Bell
Cornell University, Ithaca, New York, USA

1. Introduction
2. Team Types and Task Characteristics
3. Team Composition
4. Team Development
5. Team Processes and Performance
6. Building and Managing Effective Teams
7. Conclusion
   Further Reading

GLOSSARY

team composition The collective attributes of team members, with an emphasis on how the similarities, differences, and combination of member attributes influence team processes and effectiveness.

team development The progression of a newly formed team from a loose collection of individuals to a well-integrated collective capable of coordinating effort to accomplish the team's task.

team effectiveness A multifaceted factor defined by the degree of the team's productive output, satisfaction of member needs, and the willingness of members to continue working together.

team processes Cognitive, affective, and behavioral mechanisms that enable team members to combine the potential of their attributes and resources to resolve constraints, coordinate effort, and achieve success.

work team A collective of two or more individuals that performs organizationally relevant tasks, shares one or more common goals, interacts, exhibits task interdependencies, manages boundaries, and is embedded in a broader organizational context.

Work teams are composed of two or more individuals; perform organizationally relevant tasks; share one or more common goals; interact socially; exhibit interdependencies in task work flows, goals, and/or outcomes; maintain and manage boundaries; and are embedded in a broader organizational context that sets boundaries, constrains the team, and influences exchanges with other units in the organization. Work team effectiveness is enabled by team processes that combine individual efforts into a collective product.

1. INTRODUCTION

The idea of people working together in teams is certainly not new. It is quite likely that our human ancestors combined their skills and strength for the basic survival functions of hunting and defense, and human history is represented by groups working together to explore, achieve, and conquer. Yet, as the modern concept of work in large organizations developed during the 20th century, it was primarily centered on individual jobs. During the past two decades, however, strategic, economic, and technological forces have prompted organizations world-wide to substantially restructure work around interdependent teams. Teams allow more diverse skills and experience to bear on a problem, and they provide the potential for more flexible and adaptive responses to the unexpected. For example, organizations can use cross-functional teams to combine the expertise needed to solve complex problems and can

©2004 Elsevier Inc.
All rights reserved.

use virtual team arrangements to link these experts regardless of where they are located. Although work teams offer organizations many potential benefits, they also present challenges: Grouping people together into teams does not guarantee that they will be effective.

Most theoretical frameworks for understanding team effectiveness—either implicitly or explicitly—follow the input → process → output logic posed by McGrath in 1964. Inputs represent the characteristics of individual members (e.g., abilities, skills, personalities, and demographics), task requirements and interdependencies (i.e., work flow, goal, or outcome linkages), and organizational factors (e.g., leadership, training, and resources). From a normative perspective, these factors can be viewed as a set of resources that can contribute to team effectiveness and constraints that have to be resolved for the team to be effective. Processes represent the psychological mechanisms that enable team members to combine the potential of their characteristics and resources to resolve the constraints and achieve success. In particular, the task as a source of interdependence requirements is a key constraint that has to be resolved. Outcomes represent internal and external evaluations of team performance and the effect of the experience on team members. When team members synergistically combine the potential of their characteristics and resources to enact team processes that fit or resolve the constraints—creating synergy or process gains—their performance is effective. When team members fail to combine their potential, the resulting process loss impedes effectiveness. Indeed, the inability to capitalize on the potential of team members in cohesive decision-making groups can lead to catastrophic failures, such as those described by Janis in his 1982 book on the phenomenon he labeled groupthink.

This article discusses key factors that are responsible for effective work teams, including team types and task characteristics, member composition, team development, and processes that enable individuals to successfully combine their efforts to achieve team products.

2. TEAM TYPES AND TASK CHARACTERISTICS

One challenge to understanding team effectiveness is that teams have a variety of different forms, with new forms being invented to meet emerging needs (e.g., transnational and virtual teams). Team forms have different constraints and therefore require different process mechanisms for effective performance: One size does not fit all. In an effort to better document these differences, team researchers have developed typologies to classify team forms as shown in Table I. Other researchers have argued that classification per se does not advance understanding of the constraining factors that distinguish teams, and that a more useful approach is to identify the underlying dimensions that distinguish team types. In a 2003 review, Kozlowski and Bell posited the following set of dimensions for characterizing team constraints: (i) external environment or organizational context in terms of its dynamics and degree of required coupling; (ii) team boundary permeability and spanning; (iii) member diversity and collocation/spatial distribution; (iv) internal coupling requirements; (v) work flow interdependence with its implications for goal, role, process, and performance demands; and (vi) temporal characteristics that determine the nature of performance episodes and cycles and the team life cycle.

The team task is a critical characteristic distinguishing teams both within and across types. In 1972, Steiner presented a classic typology of team tasks based on the complexity of member interdependence and resulting coordination demands required for the combination of member effort. For example, simple additive tasks sum individual work effort or products (e.g., a rope pull or a typing pool). Conjunctive tasks are limited by the performance of the weakest team member (e.g., mountaineering), whereas performance on disjunctive tasks is determined by the strongest member (e.g., problem solving). In organizations, more complex discretionary tasks allow wide latitude for how members combine their efforts, necessitating that members monitor and coordinate effort to accomplish collective outcomes. As discussed later, it is these demands for active, coordinated interdependence that place similarly complex demands on enabling team processes. Configuring team processes to fit the demands and constraints imposed by the team's task interdependence is the key to creating effective work teams.

3. TEAM COMPOSITION

As mentioned previously, one set of inputs that influence team effectiveness is the characteristics of individual members. Recognizing that team processes and performance will be influenced by the number and type of people who are its members, considerable research has focused on the effects of team composition. One issue that

TABLE I
Types of Work Teams

Team category	*Key distinguishing characteristics*	*Examples*
Production	Core employees who cyclically produce tangible products and vary with regard to discretion from supervisor-led to semiautonomous	Automobile assembly Mining units Wood harvesting units
Service	Engage in repeated interactions with customers who have different needs, making the nature of the transactions variable	Insurance service groups Retail store teams Hospital service teams
Management	Senior managers of meaningful business units with primary responsibility for directing and coordinating lower level units under their authority	Bank management teams Health care teams Division management teams
Project	Temporary entities that execute specialized, time-constrained tasks and then disband; are generally cross-functional	New product teams R&D project groups Process improvement teams
Action/performing	Composed of interdependent experts who engage in complex, time-constrained performance events involving audiences, adversaries, or challenging environments	Military tank crews Rescue units Professional musician groups
Advisory	Temporary groups designed to solve problems and recommend solutions; work outside of, and in parallel with, production processes	Quality circles Employee involvement teams Selection committees
Crews	Characterized by the capability and necessity to form and be immediately prepared to perform together effectively; membership is dynamic and highly expert	Aircraft crews Military combat units Surgical teams
Top management	Teams composed of individuals at the executive level of an organization or corporation	Corporate executive groups Hospital top management teams
Transnational	Global or international teams; membership often spans national, cultural, and organizational boundaries	Global sales teams Multinational R&D teams
Virtual	Teams whose members are dispersed in space and interact through communication and information-sharing technologies	Distributed work teams Air traffic control teams Mission control teams Military command and control

has garnered attention concerns the best size for various types of teams. Although it is tempting to try to simplify this issue through broad-based recommendations (e.g., more is better), it appears that the "optimal" group size depends on a number of specific contingencies, such as the level of interdependency required by the team's task and the stability of the external environment.

The extent to which team processes and outcomes are influenced by the homogeneity or heterogeneity of member demographic characteristics is another issue

that has been the focus of considerable attention. Like team size, whether diversity is advantageous or detrimental to team functioning and performance appears to depend on a number of factors. Research suggests that diversity is particularly valuable when teams face creative and intellective tasks and the diversity is in terms of skills and expertise rather than demographic characteristics. Research on demographic diversity has been extended to consider the team composition effects of factors such as personality and cognitive ability on team effectiveness. Both team-level conscientiousness, a personality trait, and team cognitive ability appear to be fairly potent, positive predictors of team effectiveness. Research also suggests that other personality factors, such as agreeableness, may be important and that the effects of both personality and cognitive ability depend on the task and the amount of team member interaction required for effective performance. Future research will shed more light on such issues.

Team composition has been a popular topic because of its theoretical and practical implications. Theoretically, team composition research goes to the heart of understanding how individual attributes combine to form effective interdependent groups. An even better understanding of such issues should emerge in the near future as researchers focus more attention on clearly articulating the processes by which individual actions contribute to collective team outcomes. Practically, an understanding of team composition can serve as a valuable tool for selecting and constructing effective teams. Recommendations generated from this research can help guide staffing efforts designed to produce the optimal blend of employee characteristics and can be used to effectively manage the diversity that exists in today's transnational and virtual teams.

4. TEAM DEVELOPMENT

Existing teams with an ongoing life span will have outflows and inflows of new members, necessitating a process of socializing newcomers to adopt team norms, values, and goal expectations. Although there is a solid research foundation for understanding socialization to organizations, there is virtually no research pertaining to team socialization. In other instances, teams are formed anew. This is a common practice in organizations as project or virtual teams are formed to address specific problems. New teams go through a developmental process as individuals endeavor to enact norms, values, and goal expectations to guide their interaction and efforts.

Much of the conceptual work in this area takes a descriptive perspective in that it endeavors to characterize the natural process of group development. Stage models, such as the much imitated classic posed by Tuckman in 1965, describe a linear series of developmental stages to capture the team life cycle: forming, storming, norming, and performing. During the forming stage, team members first come together and begin to explore the group. They attempt to define the group task and to structure how they will accomplish it. There can be considerable ambiguity and differences of opinion as members realize that accomplishing the group task will be difficult. Lack of structure and ambiguity promote conflict, which typifies a shift to the storming stage. Disagreement and dismay ensue as members argue about group actions and form factions. With time, the group shifts to norming, as ground rules, roles, and goals are clarified and established. Conflict is replaced by cooperation, and members develop a sense of cohesion. With these normative expectations in place, the group shifts to the performing stage, with a focus on task accomplishment. Members are able to prevent or resolve intragroup conflict when it arises. They become attached to the team and derive satisfaction from progress toward collective goals.

Although some developmental process is necessary for members of new teams to develop a normative structure to guide interactions, this stage model illustrates that the process can be costly in terms of time and emotion. As a result, a variety of "team-building" interventions have been proposed in an effort to improve the normative structure of teams. Unfortunately, often such interventions are provided after teams have developed and enduring problems have been created.

The model proposed by Tuckman was based on groups working on unstructured tasks (i.e., clinical, therapy, and T-groups) and therefore heavily emphasizes interpersonal processes as individuals struggle to create some structure to guide their interactions. In a classic 1988 article, Gersick presented another descriptive model of group development based on the observation of 8 work and 8 student project teams that described the process as a two-stage punctuated equilibrium. The key factor in this model is the entrainment of group development to an external deadline that paced progress. Very early group interactions set norms that patterned activity to the midpoint of the groups' life cycle (deadline). At that point, there was a significant transformation—the punctuated equilibrium. Groups reorganized and focused on task achievement. Although some scholars regard the models as conflicting, Chang and colleagues

reported complementary processes of linear progression, indicative of Tuckman's and other stage models, and temporal entrainment, indicative of Gersick's punctuated equilibrium. Thus, both models exhibit evidence for their descriptive validity with respect to developmental processes that occur naturally. From an application perspective, however, the focus is more on prompting optimal processes rather than on what happens naturally. Thus, in contrast to the descriptive approach, prescriptive or normative approaches specify desirable conditions that should be created to foster better or faster development by targeting team processes that underlie team effectiveness. As discussed later, normative approaches provide a basis for designing interventions.

5. TEAM PROCESSES AND PERFORMANCE

Considerable effort has been focused on trying to understand the nature of team processes and how team processes contribute to or inhibit effective performance. There are literally thousands of research articles addressing these issues. Social psychology research on small groups has tended to focus on interpersonal processes that influence the nature and quality of member attraction and interaction, with relatively little attention to the group task. In contrast, research on work teams accepts the importance of interpersonal processes but treats the team task as central and focuses more attention on processes that reflect task-driven interactions. In their 2003 review on work teams, Kozlowski and Bell classified team processes into cognitive, affective–motivational, and behavioral mechanisms in an effort to organize this research. Descriptions of key process mechanisms that influence the performance of work teams are summarized in Table II.

Cognition is a property of individuals, but team members need to act in concert. Cognitive mechanisms attempt to capture in a collective fashion team members' task-relevant perceptions, knowledge, or information. A prominent approach posits that team processes and performance are enhanced when members share a common understanding of the task environment (mental models), its goal–role–strategy requirements (coherence), and perceptions of imperatives from the broader organizational context (climate). Much of this work has been performed within the past decade, but it has been supportive and shows promise. Other approaches posit that knowledge and information do not need to be shared per se; rather, it is the availability and access to the informational resources that make it useful to the team. For example, transactive memory is based on the idea that different team members will attend to different information and will know different things. As long as members know "who knows what," the information can be available to the team collectively. Similarly, team learning is a complex process that entails both individual and collective components. Work on these latter mechanisms is also recent, and also promising, although the empirical findings are less well established. It is likely that all these concepts will become more mature as researchers continue to shed light on cognitive processes in work teams.

Affective and motivational processes are also important. In addition to what team members know, the extent to which they bond to the team and its task, interact smoothly with minimal conflict, and are motivated to accomplish the team mission positively influences team effectiveness. For example, research indicates that task cohesion contributes positively to group task performance, whereas social cohesion (member attraction) appears to be less important to performance. On the other hand, social cohesion contributes to satisfaction and can help keep a good team viable over the long term. Research on conflict suggests that interpersonal conflict inhibits effectiveness, whereas task conflict has the potential to contribute to effectiveness when the group task is nonroutine. For example, differences of opinion in a group problem-solving task may surface a superior set of options. Work in this area is still in the process of better mapping relationships. Of particular importance is maturing work that gets at team motivational processes in the form of collective efficacy—a shared sense of group competence—that influences the goals that teams set, the effort they expend for achievement, and their persistence in the face of challenges. This line of work shows consistent contributions of collective efficacy to team effectiveness. Indeed, meta-analytic findings indicate that collective efficacy is an important predictor of team performance.

Cognitive and affective–motivational processes capture what team members think, how they feel, and what they are prepared to do. Behavioral processes represent what team members actually do to combine individual effort and action to accomplish team outcomes. One of the challenges of working in teams is that individuals have to allocate effort to accomplish their own responsibilities but also have to cooperatively aid others and the team as a whole. For simple tasks, the degree of cooperation is discretionary and

TABLE II

Cognitive, Behavioral, and Affective/Motivational Processes That Underline Team Effectiveness

Cognitive		*Affective/motivational*		*Behavioral*	
Process	*Description*	*Process*	*Description*	*Process*	*Description*
Mental models	Team members' shared, organized understanding and knowledge about key elements of the team's task environment	Cohesion	Combination of shared commitment or attraction to the team task or goal and team members' attraction to or liking of the group	Coordination	Activities required to manage interdependencies with the team work flow
Team climate	Team-level shared perceptions of important contextual factors that affect team functioning	Collective mood	Team emotion or affective tone; examines how individual-level emotions combine at the team level	Cooperation	The willful contribution of personal efforts to the completion of interdependent jobs
Team coherence	Shared comprehension of the task situation and corresponding goals, strategies, and role linkages	Collective efficacy	A team's shared belief in its own collective ability to organize and execute a course of action required to produce given levels of attainment	Communication	Taskwork communication involves exchanging task-related information and developing team solutions; teamwork communication focuses on establishing patterns of interaction
Transactive memory	Team-level shared system for encoding, storing, and retrieving information	Group potency	The collective belief of a group that it can be effective		
Team learning	Relatively permanent changes in the knowledge of an interdependent set of individuals associated with experience	Conflict (divisiveness)	Can manifest as task conflict (disagreement about task content) and/or interpersonal conflict (interpersonal incompatibilities)		

largely guided by normative expectations. Complex tasks, however, have interdependence demands that place exacting requirements on whom needs to do what and when. These demands for temporal pacing, synchronicity, and entrainment distinguish coordination from discretionary cooperation. Communication helps to enable both cooperation and coordination.

6. BUILDING AND MANAGING EFFECTIVE TEAMS

Given the increasing use and importance of work teams, one might imagine that there is substantial interest in designing, selecting, training, and leading teams to be effective. Indeed, there is. This is a huge area for application. It is also the case, however, that much of what passes for practice in this area is not based on a solid scientific foundation. It is not possible to conduct a thorough summary of this area in the brief space available. However, those areas of practice that have merit and those that are open to question can be identified. Regarding specific interventions available in the marketplace, caveat emptor is appropriate advice.

Teams are often constructed and assembled with little thought about what factors need to be in place to help ensure that they have the potential to be effective. A model proposed by Hackman in 1987 takes a prescriptive perspective on team design that specifies factors necessary to support team processes that contribute to team effectiveness. The model proposes that there must be (i) a supportive organizational context that provides necessary skills via training, information, and motivating rewards; (ii) an appropriate group structure with a team composition of the correct mix of knowledge and skills, norms to guide processes, and motivation to achieve; and (iii) coaching, leadership, and support systems that provide needed resources and reduce obstacles. Teams meeting these design specification are expected to be more effective in terms of performance, member need satisfaction, and team viability over time. This is a very useful conceptual framework for thinking about what conditions have to be created by design to promote team effectiveness.

With respect to more specific interventions, several of the areas summarized in this article have relevance. For example, team composition research has the potential to guide the combination of individual characteristics (e.g., ability, personality, knowledge, and skill) to meet team design specifications. Research on these characteristics and their influence on team performance is still in its infancy. The area of team development also has the potential to contribute to team effectiveness by creating a normative structure to guide team processes. Indeed, a variety of interventions that fall under the rubric of team-building are purported to do just that. Team-building represents a huge area of practice, but one that is based largely on anecdotal support. Meta-analytic findings indicate that there is no empirical support for an effect of team-building on objective indicators of team effectiveness. The lack of supportive findings could be due to the infrequent evaluation of team-building interventions and the difficulty in measuring improvements in team effectiveness. It could also be due to the fact that team-building interventions are most often delivered to teams well after team development has concluded. It is quite likely that it is far more difficult to change team norms after they have been established than to influence and shape them during early team development. Thus, the timing of team-building interventions may be critical to their potential to be effective. Recent theory and research point to team leaders as playing a pivotal role in the development of effective teams by instilling a shared vision of the team's mission, creating a climate supportive of that mission, instructing coherent goal–role–strategy linkages (e.g., mental models), building task cohesion, and prompting collective efficacy during early team formation and development. This work is promising.

In addition, team training is an area that has amassed a sufficient research foundation—a science of team training—to usefully guide application. For example, a 1997 review by Salas and Cannon-Bowers identifies several training techniques and targets that have been shown to improve team effectiveness. Specific techniques that have received research support include cross-training (to allow members to understand others task requirements), coordination training (to enable better combination of effort), and team leader training (to prompt team development, shared mental models, and collective efficacy). These and other team training techniques target the attitudes (e.g., collective orientation), behaviors (e.g., compensatory behavior), and cognitions (e.g., shared knowledge) that underlie team effectiveness.

7. CONCLUSION

A host of strategic, economic, and technological forces have prompted a steady shift from work organized around individuals to team-based work structures. As teams pervade modern work organizations, team effectiveness

becomes an increasingly important driver of organizational success. Fortunately, there has been commensurate growth in the amount of research devoted to understanding teams and team effectiveness in work settings and, as this article highlights, much has been learned about the many different factors that impinge upon team effectiveness. Much work still needs to be done, but existing knowledge and recent advances will spur future developments in the understanding of work teams.

See Also the Following Articles

Groups, Productivity within ■ Groupthink ■ Intergroup Relations and Culture

Further Reading

Bell, B. S., & Kozlowski, S. W. J. (2002). Virtual teams: Implications for leadership. *Group and Organization Management*, *27*, 12–49.

Chang, A., Bordia, P., & Duck, J. (2003). Punctuated equilibrium and linear progression: Toward a new understanding of group development. *Academy of Management Journal*, *46*, 106–117.

Cohen, S. G., & Bailey, D. E. (1997). What makes teams work: Group effectiveness research from the shop floor to the executive suite. *Journal of Management*, *23*, 239–290.

Gersick, C. J. G. (1988). Time and transition in work teams: Toward a new model of group development. *Academy of Management Journal*, *31*, 9–41.

Hackman, J. R. (1987). The design of work teams. In J. Lorsch (Ed.), *Handbook of organizational behavior* (pp. 315–342). New York: Prentice Hall.

Janis, I. (1982). *Groupthink: Psychological studies of policy decisions and fiascoes* (2nd ed.). Boston: Houghton Mifflin.

Kozlowski, S. W. J., & Bell, B. S. (2003). Work groups and teams in organizations. In W. C. Borman, D. R. Ilgen, & R. J. Klimoski (Eds.), *Handbook of psychology: Industrial and organizational psychology* (Vol. 12, pp. 333–375). London: Wiley.

McGrath, J. E. (1964). *Social psychology: A brief introduction*. New York: Holt, Rinehart & Winston.

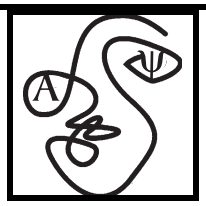
Salas, E., & Cannon-Bowers, J. A. (1997). Methods, tools, and strategies for team training. In M. A. Quinones, & A. Ehrenstein (Eds.), *Training for a rapidly changing workplace: Applications of psychological research* (pp. 249–279). Washington, DC: American Psychological Association.

Steiner, I. D. (1972). *Group process and productivity*. New York: Academic Press.

Sundstrom, E., McIntyre, M., Halfhill, T., & Richards, H. (2000). Work groups from the Hawthorne studies to work teams of the 1990s and beyond. *Group Dynamics: Theory, Research, and Practice*, *4*, 44–67.

Tuckman, B. W. (1965). Developmental sequence in small groups. *Psychological Bulletin*, *63*, 384–399.

West, M. A. (2001). The human team: Basic motivations and innovation. In N. Anderson, D. Ones, H. K. Sinangil, & C. Viswesvaran (Eds.), *Handbook of industrial, work and organizational psychology* (Vol. 2, pp. 270–288). Thousand Oaks, CA: Sage.



Writing, Teaching of

Virginia W. Berninger
University of Washington, Seattle, Washington, USA

1. Historical Perspective
2. Writing with and without Computers
3. Related Processes and Developmental Stages
4. Teaching Writing across the Grades
5. Individual Differences in Student Writers
6. Role of Teacher Knowledge
Further Reading

GLOSSARY

alphabetic principle Set of correspondences between a phoneme and the one- or two-letter spelling units that represent it that are applied in a different direction in spelling (from phoneme to spelling unit) than in reading (from spelling unit to phoneme).

automaticity Retrieval and production of a letter form or written word that is nonstrategic, effortless, direct, and rapid.

expository (essay) writing Writing that is informational (descriptive or compare or contrast) and/or argumentative (opinion or persuasive).

morphemes Word parts that, alone or attached to other word parts, convey meaning and sometimes tense, number, comparison (inflectional suffixes), or part of speech (derivational suffixes).

morphological awareness Conscious reflection and manipulation of morphemes in short-term and working memory.

narrative writing Creating a story schema with a plot or sequence of events, characters, a setting, conflict, and a resolution to the conflict or outcome.

orthographic awareness Consciously reflecting about and manipulating letter units in words in short-term and working memory that correspond to phonemes and/or morphemes.

phoneme Smallest unit of sound that (a) has no meaning of its own but makes a difference in meaning and (b) corresponds to one or more spelling units.

phonological awareness Conscious reflection and manipulation of word parts in spoken words, including phonemes, syllables, and rimes in short-term and working memory.

process writing An instructional approach that emphasizes multiple drafts and revisions rather than a one-time product.

scaffolding Teacher-provided guided assistance and feedback.

spelling unit (grapheme) A single- or multiple-letter unit in a written word that corresponds to a phoneme.

strategies Reflective approaches that are nonautomatic and employ a set of procedures to reach writing goals.

syntactic awareness Consciously reflecting about and manipulating the grammatical structures for combining words in sentences or conversational utterances in short-term and working memory.

think-alouds A research or teaching tool in which the writer or writing instructor says what he or she is thinking out loud while engaged in each step of a writing process.

Effective teaching of writing initially focuses on transcription skills (handwriting and spelling) and generating short texts and subsequently focuses on specific levels of text generation (word choice, sentence syntax, and paragraph and discourse organization), specific kinds of discourse structures (narrative and expository [informational and persuasive]), specific

©2004 Elsevier Inc.
All rights reserved.

kinds of school writing assignments (report writing, lecture notes, study skills, and creative writing), and the cognitive processes of writing (planning, translating, and reviewing/revising). The most effective teaching approaches depend on the writer's stage of writing development. Some subprocesses of writing develop at uneven rates, creating intraindividual and interindividual differences that pose challenges for teaching writing to students with diverse writing abilities. Writing is a complex process that is learned through instruction and practice over a long period of time, beginning during the preschool years and extending into the adult years.

1. HISTORICAL PERSPECTIVE

Four paradigm shifts during the final two decades of the 20th century dramatically altered writing instruction at the beginning of the 21st century. First, the notion that reading should be taught during the elementary school years and writing should be taught during the middle and high school years has been replaced with the research-based understanding that, because writing emerges during early childhood, kindergarten and primary grade school children benefit from writing instruction as well as from reading instruction.

Second, the focus shifted from pedagogy alone (the teacher's instructional behavior) to learning (the student's cognitive processes during writing). "Think-alouds" used as a research tool revealed that composing draws on the following cognitive processes: idea generating; setting goals related to content, organization, and audience; making plans to reach goals; text generating; monitoring produced text and plans in progress; and revising to repair and improve text. Teachers developed methods to teach these processes by (a) thinking aloud as they modeled a writing process (e.g., planning, drafting, reviewing), (b) providing multiple-step algorithms for children to apply in their independent writing (e.g., a checklist with hints for reviewing and revising text), (c) props such as "think–plan" sheets for helping students to plan before writing and "think–revise" sheets for helping students to revise after writing a draft, and (d) scaffolding (i.e., guided assistance and feedback) as students attempted to use the various processes in their own writing.

Third, based on insights from expert professional writers, the "process writing" movement replaced the prior emphasis on the writing product that was handed in for a grade but was seldom revised based on teacher feedback to improve the writing. Professional writers write and rewrite many drafts in which they incorporate feedback from their editors before publishing their final drafts. The process writing movement places more value on the process leading to the final product than on the initial writing product.

Fourth, the process writing movement discovered the social nature of writing and refuted the prior view that writing is an independent activity produced for an audience of one—the teacher. Students do meet individually with teachers to discuss their writing in progress, but they also receive feedback from peers and often "publish" their final versions for their peers to read and discuss.

2. WRITING WITH AND WITHOUT COMPUTERS

Students in the 21st century need to learn to write by hand as well as by keyboard. Writing with pencil or pen is still important because forming letters by hand engages the attention system in a way that a key press does not, and the enhanced attention to letter forms facilitates learning to read and spell words. Also, written communication is often needed in contexts where computer technology is not available. Finding letters on the keyboard is a necessary first step to learning to operate a word processing program through a keyboard and should be taught as explicitly as is handwriting. Computerized spell checks do not eliminate the need for systematic spelling instruction. Spell checks merely provide a check for typing errors; unless the writer can recognize the correct spelling among the choices offered, the spell check does not fix the spelling. Some writing goals are achieved best with pen or pencil, whereas computers have added value for writing goals such as revision and completion of long-term writing assignments. Voice recognition software for dictating written text has many technical difficulties to overcome. Students still need to be taught to compose written text using pencils, pens, and keyboards.

3. RELATED PROCESSES AND DEVELOPMENTAL STAGES

Writing draws on the sensory, motor, language, cognitive, memory, and executive functions of the brain, but the relative contributions of these various processes change at different developmental stages as the

brain constructs, through instructional experiences, an evolving functional writing system. Low-level skills such as handwriting and spelling are separable from high-level skills involved in composing. Low-level skills typically exert their greatest constraints early in writing development. Once low-level skills are automatic, high-level skills exert the greatest constraints on written composition. Handwriting, spelling, and composing all have their own developmental trajectories (Table I).

3.1. Handwriting Development

The normal developmental trajectory for manuscript letters, cursive letters, and keyboarding is from accurate to automatic letter production (Table I). These stages are first achieved for printing (manuscript) during the first two grades and then for cursive during the third and fourth grades. Cursive is more difficult because not only letter forms but also connecting loops must be mastered.

3.2. Spelling Development

Table I summarizes the typical developmental patterns from phonetic spelling, to phonemic spelling (alphabetic principle), to morphological spelling as well as the shift from sound-based (phonetic or alphabetic) strategies to automatic retrieval of orthographic word forms. The third and fourth stages overlap during spelling development. Also, there are tremendous individual differences in the rates at which children proceed through these stages and master them.

3.3. Composition Development

Table I summarizes the early stages of composition development. Paragraph structure with a true topic sentence and supporting propositions often does not develop until the upper elementary school grades. During the upper elementary and middle school grades, narrative writing becomes more elaborated with complex plots and character development, and proficiency with a variety of expository structures is still evolving. For the informational essay structure, proficiency with descriptive essays develops earlier and faster than does proficiency with compare-and-contrast essays. For the argument essay structure, proficiency with opinion essays develops earlier and faster than does proficiency with persuasive essays in which arguments and counterarguments are presented and the writer takes a position and defends it. Composing is mastered only after many years of instruction and practice.

3.4. Integration with Other Language Systems

Writing development is not complete until students integrate writing with the other language systems: for listening (e.g., early spelling development, taking notes during lectures), for speaking (e.g., giving speeches with audiovisual aids), and for reading (e.g., taking notes from source materials to write reports or study for tests). All of these cross-language integration activities place large demands on the executive functions of the brain for coordinating multiple processes.

4. TEACHING WRITING ACROSS THE GRADES

4.1. Kindergarten

To develop the fine motor skills for using a pencil, children should play with pegboards and blocks, string beads, use crayons and markers for free drawing, and solve paper-and-pencil mazes. To develop the strength for using a pencil, children should play with clay. Children should be taught to write and name alphabet letters accurately. Instructional activities for accomplishing these goals include the following: (a) tracing over dotted lines to form alphabet letters, (b) studying numbered arrow cues to learn the sequential component strokes in forming alphabet letters, (c) naming letters while writing them, and (d) learning special names and stories for each letter (e.g., "Mr. K" is a kicker who kicks footballs and soccer balls). Teachers should model composing short texts and teach the following strategy: "What I Think, I Can Say, What I Can Say, I Can Write." Children should be taught strategies for using cue cards (with pictures of key words with the sound for each alphabet letter or letter group) to generate the letters that spell the sounds in spoken words. Daily opportunities to compose freely and illustrate the writing with drawing should be provided.

4.2. Primary School Grades

To achieve handwriting automaticity (i.e., effortless printing of legible manuscript letters), which transfers to improved written composition, children should be asked to write letters from memory and

TABLE I
Developmental Trajectories of Component Writing Skills

Handwriting development	
First stage	Accurate (legible) production of manuscript letters Individual letters are recognizable even if not in word context
Second stage	Automatic (e.g., fast, effortless) legible production of manuscript letters
Third stage	Accurate (legible) production of cursive letters and their connecting strokes Individual letters are recognizable even if not in word context
Fourth stage	Automatic (e.g., fast, effortless) legible production of cursive letters and their connecting strokes
Keyboarding development	
First stage	Finding keyboard letters by trial and error
Second stage	Finding keyboard letters accurately and quickly with visual cues
Third stage	Finding keyboard letters accurately and automatically (e.g., fast, effortlessly) without visual cues
Spelling development	
First stage	Words spelled as perceived in speech (e.g., phonetic cues) and in letter names (e.g., /cac/, /yd/, /mt/).
Second stage	Words spelled with the alphabetic principle (i.e., correspondences between phonemes and one- and two-letter spelling units) and both plausible nonconventional (e.g., /caik/, /whide/, /emptee/) and conventional (e.g., cake, wide, empty) spellings.
Third stage	Conventional spelling of words of increasing length; of words in which partial spelling units must be memorized (due to silent letters, alternations of the alphabetic principle for spelling the same sound, and schwas such as m<u>o</u>ther, <u>u</u>pon); words with morphemic units (i.e., inflectional suffixes for tense, plurality, and comparison as well as derviational suffixes that mark part of speech); words spelled with morphological rules (e.g., for adding suffixes to the root or stem)
Fourth stage	Automatic retrieval of the whole written word from long-term memory
Composition development	
Precomposing free association	Write a comment about a word in the previous sentence that is not the topic. [Mary laughed. Laugh don't cry.]
First stage	Topic–comment: Generate a comment about the topic in the prior sentence, as in oral discourse. [Mary laughed. Tom also laughed.]
Second stage	Wheel structure representing a topic (i.e., the central hub) with multiple comments (i.e., spokes) about the topic. [I like to play tag. I like to play hide-and-seek. I like to play Monopoly.]
Third stage	Elaborated wheel structures (i.e., comments about spokes). [I like to play games. I like to play tag. I play tag with my cousins. I like to play hide-and-seek. I play that with Jim and Jane next door.]
Fourth stage	Emerging discourse structure: Simple narratives with event sequences and some mention of characters and/or setting are produced. Also, simple essays (expository texts) are produced, but in the form of knowledge telling or listing that does not always adapt the message for the audience.
Fifth stage	Refining discourse structures and emerging knowledge-transforming strategies that adapt writing to the audience. [There are many kinds of games children can play. Some are played outside. Tag and hide-and-seek are examples of outdoor games. Some are played inside. "Go fish" and Monopoly are examples of inside games.]

not merely copy them. Practicing each alphabet letter once each day—in a different random order each day—is a more efficient way in which to develop automaticity than is mindless copying of just a few letters over and over. After studying model letters with numbered arrow cues, children should cover them for a few seconds and then write them from memory. This activity should be teacher led with frequent naming of the letters because verbal names are retrieval cues that help to make the process of getting letter forms in and out of memory more automatic. Other ways in which to develop automatic letter retrieval and production include (a) writing each letter once from dictation and (b) writing letters that come before or after other letters. When cursive is introduced during third grade, the goal should initially be accurate letter form production and accurate connecting strokes.

Systematic spelling instruction for words of one and two syllables should be introduced during first grade and continued through third grade. Using high-frequency Anglo-Saxon words in the writing of primary grade school children, teachers should teach the following: (a) alphabetic principle for going from phonemes in spoken words to one- and two- letter spelling units (which contrasts with the alphabetic principle for decoding words during reading that goes from letters to phonemes), (b) alternations (the options for spelling the same phoneme), (c) predictable spelling patterns larger than alphabetic principle (e.g., word families such as /ight/), and (d) morphemic units (e.g., inflectional suffixes that have multiple sound–spelling relationships for plurals [/s/ as in cats, /s/ as in bees, and /es/ as in busses] and for tense [/ed/ as in looked, opened, and added]). The high-frequency words should be practiced alone and in sentences. Daily spelling-to-sentence dictation is more effective than weekly spelling tests with lists only. To transfer knowledge of single-word spelling to composing, children should be asked to write compositions containing a set of high-frequency words and to keep frequently misspelled words in their compositions in a personal dictionary.

Free journal writing is not sufficient for written composition. Children should be given compositional prompts (i.e., topics) to write about, compose, and share with peers daily. Teachers should give daily feedback and help children to set goals for the next day (e.g., to write one more sentence, to choose more interesting words) based on a visible record of the progress each day (e.g., a graph of how many words or sentences they have written).

4.3. Upper Elementary School Grades

The transition from third to fourth grade is important in writing development because the writing requirements of the curriculum increase substantially during fourth grade and beyond. Students are now expected to explain in writing their understanding of passages they read and to write reports based on source materials for the content subjects. Children whose letter production and keyboarding are not automatic and whose spelling is not grade appropriate will face special challenges in keeping up with the increasing writing requirements. Their teachers need to provide individually tailored instruction in handwriting and/or spelling (and monitor progress in development of transcription skills) along with composition instruction. Explicit instruction in using word processing programs for writing assignments should be introduced for all students.

Continued systematic spelling instruction for longer, morphologically complex words of Latinate origin (often three syllables or more and containing derivational suffixes that signal grammatical information about part of speech) and of Greek origin should be provided for all students. Morphological spelling rules should be taught for adding suffixes to roots or bases (i.e., roots already having suffixes). Strategies for memorizing word-specific spellings containing schwas or homonyms should also be taught. In general, spelling benefits from training in phonological, morphological, and orthographic awareness of words, their parts, and their interrelationships.

The instructional goals for written composition at this stage of writing development draw on reading–writing connections and include (a) learning to write well-formed paragraphs and longer discourse structures for specific genre (narrative and expository [essay]), (b) learning to express through writing the author's understanding of a text written by another author (reading comprehension), and (c) learning to take notes from source material, organize notes, and write reports based on the notes for content subjects such as science and social studies as well as for language arts assignments such as book reports. Effective instructional techniques include (a) examination of model paragraphs and model texts that contain prototypical structures that children can use in their own writing, (b) summarization activities in which children write the main idea and most important supporting details for the texts they have read, and (c) teacher modeling and teaching of explicit strategies for each of the steps in writing reports. Composing instruction should aim at each level of

language involved in producing quality text: word, sentence, and discourse. Vocabulary development and word awareness activities help children to improve in quality of word choice. Three kinds of syntactic awareness activities improve quality of sentence writing: unscrambling randomly presented words to create a grammatically acceptable sentence, crossing out unnecessary words in overly wordy sentences, and using sentence-combining activities that express in one well-formed sentence the ideas contained in two other sentences. Graphic organizers and outlining can be used to teach organizational techniques at the discourse level. Checklists for reviewing and revising at each level of language are also helpful.

4.4. Middle and High School

The most important writing goal during the middle and high school developmental stage is to acquire the cognitive ability to transition from knowledge telling to knowledge transforming so as to meet the needs of the audience more effectively. Explicit instruction in executive functions and internalization of writing plans is critical for this goal. Student writers benefit greatly from (a) teacher–student dialogues aimed at engaging students intellectually in the planning, translating, and reviewing/revising processes of composing; (b) use of computers in the writing program and continuing instruction in keyboarding and word processing; and (c) daily teacher feedback about student performance displayed visibly (e.g., with graphs) and linked to setting goals for the next writing activity. The goal(s) for writing instruction should be broader than merely teaching sentence syntax and should include genre-specific discourse writing skills and strategies for writing for multiple purposes, including (a) completing long-term writing assignments that are revised in response to feedback along the way, (b) report writing for content subjects such as social studies and science, (c) creative writing to express original ideas and experiment with genres such as poetry and plays, (d) note taking during lectures, (e) note taking during textbook reading assignments, (f) using notes from lectures and textbooks to study for tests, and (g) test taking (essay writing tests). Some middle and high school students still need help with basic transcription skills (handwriting and spelling), vocabulary use and word choice, syntactic structures, and discourse structures.

Most adolescents make the transition from learning to write to writing to learn and writing to solve problems. Writing plays an important role in the transition from schooling to adult life. High school graduation is greatly dependent on completing writing assignments. Passing state standards in writing on high-stakes tests is becoming a requirement for high school graduation in some states. College entrance tests are increasingly requiring writing assessments. Many college-level courses also require advanced writing skills. Increasingly in the 21st century, writing is affecting the options open to adults for earning a living. Many jobs involve the use of computers that require both writing and reading to enter written language and view it on the monitors. Teaching the use of computers in writing instruction is a matter of teaching writing–reading connections.

5. INDIVIDUAL DIFFERENCES IN STUDENT WRITERS

At all grade levels, individual students show some normal variations in their relative strengths and weaknesses with specific writing skills and related processes. However, in some cases, specific writing skills fall outside the normal range and students have specific learning disabilities despite their normal intelligence or thinking ability. In such cases, the students should be referred to a multidisciplinary team that includes a psychologist, a speech and language pathologist, and an occupational or physical therapist. Common specific learning disabilities include dyslexia, which is impaired single-word reading and spelling, and dysgraphia, which is impaired handwriting and/or spelling without impaired reading. Some children have both dyslexia and dysgraphia. Other students have language learning disabilities and have trouble using oral language to learn in school and learning written language. Students with specific learning disabilities require specialized instruction.

6. ROLE OF TEACHER KNOWLEDGE

Teachers who have knowledge of research findings about the writing process and effective writing instruction are more likely to help their students to achieve developmentally appropriate writing goals than are those who do not. However, teachers who have research knowledge also need to evaluate the effectiveness of their implementation of research-supported instructional practices in their own classrooms for their own students. The application of research to

TABLE II
Instructional Design, Pedagogical, and Evaluation/Modification Principles

A. Instructional design principles for planning instruction at specific levels
   1. Incidental writing instruction—teaching skills only as individual students require special kinds of assistance—is not sufficient at any level. Lessons in handwriting, keyboarding, spelling, and composing, with and without computers, should be highly preplanned so that teachers can respond flexibly to individual students' needs without losing sight of the planned overall instructional goals in a systematic program of instruction.
   2. At all levels, plan how instructional components are packaged close in time within a lesson. Teaching to all levels of language close in time within a single lesson helps to overcome the temporal limitations of working memory and has been shown to create functional writing systems in which component writing processes are coordinated.
   3. At the middle and high school levels, plan how lessons are sequenced over the course of the week, month, semester, and year to teach the multiple steps involved in long-range writing assignments and complex skill development.

B. Pedagogical principles for implementing instruction at all levels
   1. Aim instruction at a level that is not too easy or too hard but is intellectually engaging and should lead to growth.
   2. Teach for transfer. That is, provide opportunity to practice skills in isolation and then in multiple contexts to promote generalization to novel contexts.
   3. Teach low-level skills to automaticity to free up resource-limited working memory for high-level goals.
   4. Avoid habituation (i.e., where the brain becomes unresponsive to repetitive activity) by teaching automatization of low-level skills only for brief periods of time.
   5. Also teach high-level, nonautomatic skills in an explicit way with teacher-prompted cueing of relevant dimensions, teacher modeling, teacher–student interactive dialogue, teacher-devised strategies, and teacher-led scaffolding of the application of all of these as students apply them to their writing.
   6. Teach executive functions for internally guiding and self-regulating the writing process.

C. Evaluation/Modification principles for monitoring student progress at all levels
   1. Monitor student progress daily, weekly, and monthly as well as at the beginning, middle, and end of the year with curriculum-based, criterion-referenced, and norm-referenced measures. Use results for instructional decision making (revise if necessary just as in writing).
   2. Provide visible records (e.g., through graphs) of student performance for teachers and students to review on a regular basis.
   3. Adjust teaching approach if necessary if students progress too slowly or quickly.

Source. Adapted from Wong and Berninger (2004).

practice is not always simple and straightforward, and specific implementations need to be evaluated. Like writers who must plan, translate, and review and revise their writing, teachers must plan, translate, and review and revise their applications of research to practice. Table II summarizes research-supported general principles that are helpful in this regard.

See Also the Following Articles

Learning ■ Literacy, Improvement of ■ Reading, Teaching of

Further Reading

Alley, G., & Deschler, D. (1979). *Teaching the learning-disabled adolescent: Strategies and methods*. Denver, CO: Love Publishing.

Berninger, V., & Abbott, S. (2003). *PAL research-supported reading and writing lessons*. San Antonio, TX: Psychological Corporation.

Berninger, V., & Amtmann, D. (2003). Preventing written expression disabilities through early and continuing assessment and intervention for handwriting and/or spelling problems: Research into practice. In H. L. Swanson, K. Harris, & S. Graham (Eds.), *Handbook of research on learning disabilities* (pp. 345–363). New York: Guilford.

Berninger, V., & Richards, T. (2002). *Brain literacy for educators and psychologists*. San Diego: Academic Press/Elsevier.

Clay, M. (1982). Research update: Leaning and teaching writing—A developmental perspective. *Language Arts*, *59*, 65–70.

Englert, C. S., Raphael, T., Anderson, L., Anthony, H., & Stevens, D. (1991). Making strategies and self-talk visible: Writing instruction in regular and special education classrooms. *American Educational Research Journal*, *28*, 337–372.

Graves, D. (1983). *Writing: Teachers and children at work*. Exeter, NH: Heinemann.

Harris, K., & Graham, S. (1996). *Making the writing process work: Strategies for composition and self-regulation*. Cambridge, MA: Brookline.

Hayes, J. R., & Flower, L. S. (1980). Identifying the organization of writing processes. In L. W. Gregg, & E. R. Steinberg (Eds.), *Cognitive processes in writing* (pp. 3–30). Hillsdale, NJ: Lawrence Erlbaum.

Hillocks, G. (1984). What works in teaching composition: A meta-analysis of experimental treatment studies. *American Journal of Education*, *93*, 133–170.

Klein, P. (1999). Reopening inquiry into cognitive processes in writing-to-learn. *Educational Psychology Review*, *11*, 203–270.

MacArthur, C., & Graham, S. (1987). Learning disabled students' composing with three methods: Handwriting, dictation, and word processing. *Journal of Special Education*, *21*, 2–42.

Scardamalia, M., & Bereiter, C. (1989). Knowledge telling and knowledge transforming in written composition. In S. Rosenberg (Ed.), *Advances in applied psycholinguistics*, Vol. 2: *Reading, writing, and language learning* (pp. 142–175). Cambridge, UK: Cambridge University Press.

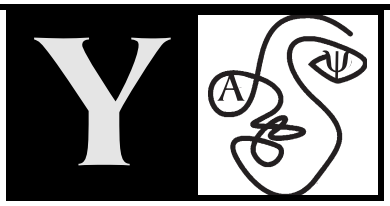
Treiman, R. (1993). *Beginning to spell*. Cambridge, UK: Cambridge University Press.

Venezky, R. (1999). *The American way of spelling*. New York: Guilford.

Wong, B., & Berninger, V. (2004). Cognitive processes of teachers in implementing composition research in elementary, middle, and high school classrooms. In B. Shulman, K. Apel, B. Ehren, E. Silliman, & A. Stone (Eds.), *Handbook of language and literacy development and disorders*. New York: Guilford.

Wong, B. Y. L., Butler, D. L., Ficzere, S. A., & Kuperis, S. (1996). Teaching adolescents with learning disabilities and low achievers to plan, write, and revise opinion essays. *Journal of Learning Disabilities*, *29*, 197–212.

Wong, B. Y. L., Butler, D. L., Ficzere, S. A., & Kuperis, S. (1997). Teaching adolescents with learning disabilities and low achievers to plan, write, and revise compare-and-contrast essays. *Learning Disabilities Research & Practice*, *12*, 2–15.



Youth, Employment of

Fred W. Vondracek

The Pennsylvania State University, University Park, Pennsylvania, USA

1. Youth Employment Defined
2. The Context of Youth Employment
3. The Developmental Significance of Youth Employment
4. Strategies for Facilitating Youth Employment
Further Reading

GLOSSARY

apprenticeship model A highly structured means for introducing youth to the world of work, for training them, and for integrating them into the labor force; it has been most successfully employed in Germany.

historical context Conditions of youth employment in different historical epochs.

occupational identity A sense of belonging and commitment to an occupational group, typically acquired after a period of exploring various career options.

part-time employment Paid work of less than 40 hours per week, typically performed by individuals who are also full-time students.

youth employment Paid work performed by individuals 16–24 years old.

Youth employment is a fact of life for the vast majority of young people in the United States. The same does not necessarily hold for youth from other countries or for youth from different historical periods. Moreover, the developmental significance of youth employment varies for each individual depending on the individual's characteristics, intensity of the work experience, and family, community, economic, and cultural contexts within which the youth's work is occurring. Thus, youth employment may result in both positive and negative consequences, although when properly structured it can be a valuable experience on the way to adulthood and independence.

1. YOUTH EMPLOYMENT DEFINED

Although there is no universal agreement on the definitions of youth or employment, for present purposes, youth employment is defined as paid work by individuals 16–24 years old. Both part-time employment (usually occurring among individuals who are also full-time students) and full-time employment are included in this definition.

2. THE CONTEXT OF YOUTH EMPLOYMENT

It is impossible to discuss youth employment without taking into account the various levels of context within which it takes place. Youth work takes place in relation to the families and communities to which young people belong. It occurs in relation to the cultural norms and expectations that prevail among these families and communities, and it is defined largely by the economic circumstances that occur at any given historical period. Nevertheless, because individual

©2004 Elsevier Inc.
All rights reserved.

circumstances vary greatly from one person to another, there can be considerable variation in how youth employment is defined and experienced, even when individuals share historical time and macrocontextual circumstances.

2.1. Historical Context of Youth Employment

In preindustrial societies, children and adolescents were integral to the economic survival of families. As early as they were able to perform any useful work, they were pressed into service, working side-by-side with adults, helping to produce food and to provide shelter and safety. As industrialization progressed, parents increasingly found work away from home, largely removing them as effective socialization-to-work agents. Adolescents were increasingly faced with the necessity of discovering their own path into the world of work and acquiring their own occupational identity and with it their independence. Moreover, as industrialized nations became wealthier and their commerce became more complex, they required a more highly educated and trained workforce. As a result, education rather than labor began to occupy most children's and adolescents' time, and much of the education and training was explicitly designed to prepare youth for an ever-expanding variety of occupations.

2.2. Cultural and Socioeconomic Context of Youth Employment

As noted previously, youth employment varies not only by historical time but also by cultural and socioeconomic context. For example, in some cultures employment opportunities for males and females differ greatly. Accordingly, boys and girls are differentially socialized into early work experiences and they emerge with different expectations and aspirations about their future occupational careers. Economic necessity may also impact the nature and extent of youth employment. Thus, youth from poor socioeconomic backgrounds are more likely to quit school and go to work to contribute to their own and their family's economic survival. In contrast, youth employment is typically optional for youth from more privileged backgrounds, where it may be considered a step toward "growing up" or simply a "good experience."

3. THE DEVELOPMENTAL SIGNIFICANCE OF YOUTH EMPLOYMENT

Although a fair amount has been written about the detrimental impact of child labor on children's development, the impact of work on development in adolescence and early adulthood has not received adequate attention by researchers. This is surprising in view of the fact that most adolescents in the United States have a good deal of work experience by the time they graduate from high school. Data from the Monitoring the Future study at the University of Michigan indicate that approximately three-fourths of boys and girls work part-time while they are in high school, and of these, more than one-third work more than 20 hours per week. However, researchers appear to be deeply divided on the issue of whether youth employment has generally positive or negative consequences for development.

3.1. Negative Effects of Youth Employment

When students are employed part-time, some energy and time that might be devoted to studying is diverted, creating the potential for compromised school performance and disrupted social relationships. Moreover, young part-time workers are more likely to be exposed to (and thus learn) negative adult attitudes toward working as well as other adult behaviors, such as smoking and drinking, which are viewed as problem behaviors when exhibited by adolescents. The causal direction of these negative attitudes and behaviors is unlikely to be the same in all cases, however. Some researchers have found that poor school performance and adult-like behaviors precipitate early entrance into the labor force, whereas others have found that these behaviors follow part-time work, especially if it involves many hours per week. Contributing to this negative scenario is the fact that most youth employment opportunities occur in low-level work settings, such as fast-food establishments, or in low skill positions, such as janitor or clerk. Adolescents from low socioeconomic status families may get "stuck" in such low-level jobs and foreclose on any real opportunity to be occupationally upwardly mobile. Others may view working in these circumstances as completely unrelated to their aspirations and ultimate career destination, resulting in disinterest or cynical attitudes toward work in general.

3.2. Positive Effects of Youth Employment

Although the potential negative consequences of youth employment have received what some consider an inordinate amount of attention, most youth, as well as their families, feel overwhelmingly positive about part-time work. They tend to adhere to the idea that part-time work pursued in moderation represents a beneficial experience. Jeylan Mortimer and colleagues at the University of Minnesota have contributed significant findings regarding youth employment through their Youth Development Study. They found that stable work, limited to less than 20 hours per week and including good relationships with supervisors, tended to buffer the effects of family problems and to promote positive vocational development. Other researchers have confirmed that young people who work for pay tend to learn valuable lessons in self-discipline and that they often develop important insights into aspects of work that they value highly and those that they do not value. Moreover, youth employment is often instrumental in helping young workers to know what they do not want to do—arguably at least as important as knowing what they do want to do. Equally important, youth employment facilitates the development of a vocational or occupational identity, which is a key component of one's overall identity. Thus, in important ways, youth employment assists adolescents and young adults in answering the question, "Who am I?"

4. STRATEGIES FOR FACILITATING YOUTH EMPLOYMENT

Youth employment, when properly structured and integrated into the lives of adolescents and young adults, can be a valuable experience in the transition to adulthood and independence. Paid part-time work may be particularly valuable when it involves moderate hours and is viewed as complementary to the completion of schooling or training. In some ways, however, paid full-time work may be even more important for those who leave school before they complete it and those who go to work immediately after high school completion. For them, the alternatives to full-time work are underemployment or unemployment, neither of which are likely to promote a smooth transition into adulthood. As a result, there has been a dual approach to facilitating youth employment, with one branch placing a heavy emphasis on making valuable part-time learning experiences available for youth who are still in school and the other focusing on a smooth transition of youthful workers into the labor force, thereby avoiding the risks associated with youth unemployment.

Developed countries have often focused on integrating work experience and schooling, putting in place policies and programs that have been derived from the highly successful German apprenticeship model. Partnerships between public and private sectors and between schools and businesses have often been the foundation of such efforts. In less developed countries, the primary focus has been on providing employment for young people that is safe and offers them at least a reasonable chance of supporting their young families. It should be clear, therefore, that just as youth employment is quite variable in nature and highly context specific, so are methods to facilitate youth employment.

See Also the Following Articles

Part-Time Work ■ School-to-Work Transition

Further Reading

Barling, J., & Kelloway, E. K. (Eds.). (1999). *Young workers: Varieties of experience*. Washington, DC: American Psychological Association.

Hamilton, S. F. (1990). *Apprenticeship for adulthood: Preparing youth for the future*. New York: Free Press.

Heinz, W. (Ed.). (2000). *From education to work: Cross national perspectives*. Cambridge, UK: Cambridge University Press.

Mortimer, J. T., & Finch, M. D. (Eds.). (1996). *Adolescents, work, and family: An intergenerational developmental analysis*. Thousand Oaks, CA: Sage.

Petersen, A. C., & Mortimer, J. T. (Eds.). (1994). *Youth unemployment and society*. Cambridge, UK: Cambridge University Press.

Schulenberg, J., Maggs, J. L., & Hurrelmann, K. (Eds.). (1997). *Health risks and developmental transitions during adolescence*. Cambridge, UK: Cambridge University Press.

Stern, D., & Eichorn, D. (Eds.). (1989). *Adolescence and work: Influences of social structure, labor markets, and culture*. Hillsdale, NJ: Erlbaum.

Waddell, S. (2001). *What works in engaging business in youth employment and livelihood strategies*. Baltimore, MD: International Youth Foundation.